Among Marysville workers, there were very few employees who declined to participate in the earlier study by Lockey et al. (1984), where 512 out of 530 employees were included, but there is potential for selection bias in the follow-up by Rohs et al. (2008), where only 280 employees out of the original cohort were evaluated. Rohs et al. (2008) state that employees hired in 1973 or earlier (when exposure estimates were more uncertain) were more likely to participate compared to employees hired after 1973, and while the range of cumulative Libby Amphibole asbestos exposure was similar between participants and nonparticipants, participants did have higher mean cumulative exposure estimates. While it is accurate that exposure levels were uncertain before sampling began at Marysville in 1972, it is also accurate that exposures were much lower beginning in 1974, when additional industrial hygiene controls were implemented. Thus, persons hired ≤1973 had higher exposure (if less perfectly measured), while those hired ≥1974 had lower exposure, and likely less disease (under an assumption of an exposure-response effect). Thus, we might assume that the prevalence rates in nonparticipants are likely lower than in participants. The self-selection to participate in the study is dependent on the exposure, thus leading to dependent censoring and potential selection bias (see Section 4.1.3 for a discussion of this potential selection bias). However, Rohs et al. (2008) conducted a sensitivity analysis assuming that all living nonparticipants had no pleural changes and report a similar significant trend of increased pleural changes by exposure quartile. In contrast, participation rates for the Libby worker studies were much higher (see above), and there is no indication of potential bias in selection of these study participants (Amandus et al., 1987b; McDonald et al., 1986b).

Both studies of Libby workers also evaluated age and smoking as potential confounders of the association between Libby Amphibole asbestos exposure and radiographic abnormalities. McDonald et al. (1986b) report that both age and cumulative exposure are significant predictors of small opacities and pleural abnormalities in the study of current and former workers, providing regression coefficients for cumulative exposure, age, and smoking status. Amandus et al. (1987b) report that although cumulative exposure and age are both significant predictors for small opacities, cumulative exposure was not significantly related to pleural abnormalities when age is included in the model, thus limiting the usefulness of these data for RfC derivation based on pleural abnormalities. Neither study of Libby workers addressed gender, body mass index

(BMI), or time from first exposure, although both studies excluded workers with other asbestos/dusty trade occupations.

With respect to the Marysville, OH worker cohort, Lockey et al. (1984) only matched on age in their analysis. The follow-up examination by Rohs et al. (2008) included information on several important covariates, including age, gender, hire date, prior exposure to asbestos, BMI, and smoking history. Hire date and age were significantly associated with the prevalence of pleural abnormalities, and results are presented considering these covariates.

5.2.1.3. *Evaluation of Exposure Assessment in Candidate Studies*

For both the O.M. Scott facility in Marysville, OH and the Libby, MT facilities, exposure estimates rely primarily on fiber counts using phase contrast microscopy (PCM) and reconstruction of earlier exposures from company records, employee interviews, and the professional judgment of the researchers estimating historical exposures (Amandus et al., 1987a; McDonald et al., 1986a; Lockey et al., 1984). Work histories for the Libby worker cohort were extracted from company employment records, while work histories for the Marysville cohort were self-reported.

The two studies of workers in Libby, MT used similar exposure estimation, based on the same fiber measurements and work records (Amandus et al., 1987b; McDonald et al., 1986a). As discussed in Section 4.1.1.2, exposures prior to 1968 are not based on fiber measurements by PCM and, thus, are more uncertain that later exposure estimates.[25] The study population of McDonald et al. (1986b) included current and former workers, with 26% of participants over 60 and 40% of participants between 40–59 years of age at the time of their X-ray in 1983. Although tenure and dates of employment are not reported, exposure estimates for this study group would include the less-certain exposure estimates prior to 1968 (McDonald et al., 1986a). However, Amandus et al. (1987b) studied workers still employed during 1975–1982 (i.e., excluding those terminated prior to 1975) who had at least 5 years of employment. The average tenure of the study participants was 14 years. Although both studies have the limitation of less-certain exposure estimates prior to 1968, based on study design, the Amandus et al. (1987b)

[25] Exposures in the dry mill at Libby, MT, prior to 1967 were estimated from total dust measurements based on site—specific conversion ratios. Exposures for all other location operations prior to 1968 were estimated because no air sampling data were available (Amandus et al., 1987a; McDonald et al., 1986b).

study group includes a greater proportion of more recent workers. However, neither researcher assessed these uncertainties nor the impact of early exposure estimates on the apparent exposure-response relationship.

Another source of uncertainty in exposure estimates for this cohort is possible community/nonoccupational exposures. Members of the Libby worker cohort may have lived in Libby prior to/after employment and resided in Libby and surrounding areas during employment. In both cases, there may have been community exposures to Libby Amphibole asbestos that are not captured in occupational-based cumulative exposure metrics. This unmeasured nonoccupational exposure may be low relative to the estimated occupational exposures, but is, nevertheless, a source of uncertainty in estimating the exposure-response relationship.

The quality of the exposure assessment also changed over time in the Marysville cohort (Rohs et al., 2008; Lockey, 1985). Industrial hygiene measurements based on PCM analysis are available for the O.M. Scott facility beginning in 1972, although personal breathing zone samples were not available until 1976 (Rohs et al., 2008). Thus, exposure levels for all job tasks prior to 1972 are estimates from later sampling events. Additionally, air sampling data were not available for several job tasks until the late 1970s. For example, air-sampling data were only available for two of seven job tasks in the trionizing department beginning in 1973 (expander and dryer). All others have dates of 1976 or later [see Table 10, Lockey (1985)]. The installation of exposure control equipment in 1974 adds to the uncertainty in early exposures estimated from sampling in later years. There is uncertainty when the Libby ore was first used in the facility. Company records indicated that the date was between 1957 and 1960, and the University of Cincinnati used the best-available information from focus group interviews to assign the first usage of Libby ore in 1959 (see Appendix F).

EPA has collaborated with the University of Cincinnati research team to better evaluate historical exposures at the O.M. Scott facility in Marysville, OH (see Appendix F). Although no air-sampling results were found prior to 1972, additional information on plant processes from other records and employee interviews has resulted in updated exposure estimates (see Section 5.2.3.1). These refined estimates of the historical exposure improve exposure characterization for the Marysville worker cohort over previous publications.

5.2.1.3.1. *Evaluation of outcome assessment in candidate studies*

In all four candidate studies, outcomes were assessed using chest radiographs independently evaluated by multiple readers. However, there were differences in the standards used for evaluation of radiographic changes, as well as timing and quality of the radiographs. The two studies in Libby workers (Amandus et al., 1987b; McDonald et al., 1986b) used similar outcome-assessment procedures, with radiographs evaluated by three readers according to 1980 ILO standards. Two different sets of standards were used to evaluate radiographs in the Marysville cohort. The first study used modified 1971 ILO standards (modifications not stipulated) (Lockey et al., 1984), while the follow-up study used the updated 2000 ILO standards (Rohs et al., 2008).

Radiograph quality may also impact outcome assessment. In McDonald et al. (1986b), which used radiographs taken in 1983 specifically for the study, 7% of films were classed as "poor quality" (some technical defect impairing the pneumoconiosis classification) and 0.4% as "unreadable." Amandus et al. (1987b), which used available radiographs taken over a wide time period (1975 to 1982), report that the proportion of films rated as "poor quality" ranged from 14.7% to 22.8% depending on the reader. In the Marysville cohort, Lockey et al. (1984) state that "...radiographs that could not be interpreted because of poor quality were repeated" (p. 953). Rohs et al. (2008) do not report the percentage of films rated as "poor quality" but do note that 7 out of 298 (2.3%) radiographs taken were considered unreadable.

5.2.1.3.2. *Selection of principal cohort*

Based on the criteria set out in Table 5-2 and the above evaluation, the update of the Marysville, OH worker cohort (Rohs et al., 2008) is the preferred cohort. The main advantages of the Marysville, OH worker cohort over the two studies of pleural and lung abnormalities in the workers in Libby, MT are:

1) Adequate follow-up time and the availability of time from first exposure data for evaluation,

2) Minimal exposure to Libby Amphibole asbestos outside of the workplace,

3) Better quality radiographs, and use of the most recent ILO reading guidelines in the cohort update,

4) Data are more appropriate for low-dose extrapolation—a lower range of cumulative exposures for the study participants (n = 280), compared to Libby workers,

5) The data allow consideration of more covariates and potential confounders (e.g., BMI, smoking status, age),

6) The presence of a demonstrated exposure-response relationship for Libby amphibole asbestos exposure and radiographic abnormalities—in contrast to the study by Amandus et al. (1987b), which does not support an exposure-response relationship for pleural abnormalities based on the cumulative exposure metric (when age is included as a covariate).

The disadvantages of the Marysville, OH cohort compared to the two studies of pleural and lung abnormalities in the workers in Libby, MT are:

1) Approximately 70% of the Marysville, OH cohort were hired before 1972 when there were no measured exposure data [Rohs et al. (2008), and Lockey et al. (1984) study].

2) Participants in Rohs et al. (2008) were self-selected, with greater participation among older employees and those who began work prior to 1973 when exposures were relatively higher. This is a potential source of bias in study population selection analyzed by Rohs et al. (see Section 4.1.3).

3) Exposure estimates are based on self-reported work histories. In this case, there is some uncertainty in the employment history, and some individuals had extensive overtime work. Employment history was self-reported during interviews with each individual for the original study (i.e., Lockey et al., 1984), and errors in this process could affect assigned Libby Amphibole asbestos exposure estimates for this cohort.

5.2.1.4. *Selection of Critical Effect*

There are several endpoints that are suitable for consideration for the derivation of an RfC for Libby Amphibole asbestos where health effects data and exposure information are available in the principal study (Rohs et al., 2008; Lockey et al., 1984): (1) parenchymal changes viewed as small opacities in the lung; (2) blunting of the costophrenic angle (measured between the rib cage and the diaphragm); or (3) pleural thickening (both localized and diffuse). Each of these effects is an irreversible pathological lesion (ATS, 2004). As the available epidemiologic

studies describe these endpoints as viewed on standard X-rays (see Text Box 5-1), it is important to understand the distinction between what is viewed on the radiograph versus the underlying biologic lesion. The following discussion reviews the health effects associated with each of these radiographic abnormalities observed in workers exposed to Libby Amphibole asbestos.

Text Box 5-1. Radiographic Abnormalities of the Lung and Pleura

Parenchymal changes in the lung (small opacities): The small opacities viewed within the lung (interstitial changes) are indicative of pneumoconiosis and are associated with exposure to not only mineral fibers, but also mineral dust and silica. The radiographic signs of pneumoconiosis begin as small localized areas of scarring in the lung tissue and can progress to significant scarring and lung function deficits. The ILO standards provide a scheme for grading the severity of the small opacities; the size, shape, and profusion of the small opacities are recorded, as well as the affected zone of the lung (ILO, 2002).

Obliteration of the costophrenic angle: The costophrenic angle (CPA) is measured as the angle between the ribcage and the diaphragm on a posterior anterior-viewed radiograph (the costophrenic recess). When CPA blunting or obliteration is noted on a radiograph, it is recorded as present or absent (ILO, 2002). Obliteration of the CPA may occur in the absence of other radiographic signs.

Pleural thickening: The pleural lining around the lungs (visceral pleura) and along the chest wall and diaphragm (parietal pleura) may thicken due to fibrosis and collagen deposits. Pleural thickening (all sites) is reported as either localized pleural thickening (LPT) or diffuse pleural thickening (DPT). DPT of the chest wall may be reported as in-profile or face on, and is recorded on the lateral chest wall "only in the presence of and in continuity with, an obliterated costophrenic angle" (ILO, 2002). Localized pleural thickening may also be viewed in-profile or face-on and is generally a pleural plaque (parietal). Calcification is noted where present (ILO, 2002).

5.2.2. Evaluation of Radiographic Lesions as Potential Critical Effects

5.2.2.1. *Health Effects of Parenchymal Changes as Small Opacities Viewed on Standard Radiographs*

Radiographic evidence of small opacities in the lung is evidence of fibrotic scarring of lung tissue consistent with mineral dust and mineral fiber toxicity. The scarring of the parenchymal tissue of the lung contributes to measured changes in pulmonary function, including obstructive pulmonary deficits from narrowing airways, restrictive pulmonary deficits from impacting the elasticity of the lung as well as decrements in gas exchange. However, although data across the mineral fiber literature strongly support a finding of functional deficits where small opacities are visible on radiographs, the data also indicate that deficits in pulmonary function (consistent with interstitial fibrosis) are seen before these changes are detected by

radiographic examination. Thus, changes in lung function may occur before the fibrotic lesions can be detected on standard radiographs (ATS, 2004; Broderick et al., 1992). For example, decreased Carbon monoxide (CO) diffusion is a sign of reduced gas exchange in the pulmonary region of the lung and is observed in workers exposed to other types of asbestos even when small opacities are absent on radiographs. Similarly, obstructive deficits in lung function may be observed without radiographic signs for fibrotic lesions of small opacities. As decreased diffusion and obstructive deficits are mechanistically linked to changes in the parenchymal tissue these data suggest radiographs may not be sensitive enough to detect and protect against small localized lesions in parenchymal tissue of the lung. Radiographic evidence of small opacities indicates interstitial damage of the lung paremchyma, is associated with decreased pulmonary function and considered evidence of an adverse health effect. Thus, small opacities are an appropriate endpoint for RfC derivation. However, as there is evidence of functional changes in lung function from lesions not detectable on conventional radiographs, more sensitive endpoints should be considered.

5.2.2.2. *Health Effects of Diffuse Pleural Thickening (DPT) Viewed on Standard Radiographs*

DPT is a fibrotic lesion (often described as a basket weave of collagen) in the visceral pleura that encases each lobe of the lungs. The fibrotic lesion restricts the ability of the lung to expand mechanically, as well as by reducing the available volume (where thickening has progressed) (Jones et al., 1988) and DPT is strongly associated with reduced lung function (ATS, 2004). There are consistent reports of impaired lung function associated with DPT in asbestos-exposed populations (Broderick et al., 1992; Kilburn and Warshaw, 1991; Bourbeau et al., 1990). A cross-sectional study of men (n = 1,298) exposed to asbestos through various trades (e.g., boiler makers, welders, plumbers/pipefitters) included chest radiographs and spirometry (Kilburn and Warshaw, 1991). When considering the effect of DPT (with costophrenic angle [CPA] blunting) on radiographic function, FVC, FEV1, and FEF25-75[26] were all significantly reduced (85, 79, and 66% of predicted values, respectively) as compared with individuals with calcification or plaques only in men with no signs of small opacities (ILO

[26] Forced Vital Capacity (FVC); Forced Expiratory Volume in 1 second (FEV1) and Percent FVC (FEV%) = [(100 × FEV1) ÷ FVC, FEF25-75, is the expiratory flow between 25% and 75% of the FEV.]

profusion score of 0/0 or 0/1) ($p < 0.0001$). The relationship between pleural fibrosis and FVC was studied in asbestos-exposed sheet metal workers ($N = 1,211$) where not only the type of thickening (discrete versus diffuse) (ILO, 1980) but also CPA involvement and the location of the thickening were taken into consideration (Broderick et al., 1992). Univariate analysis indicated FVC was decreased by both DPT (with CPA blunting) and circumscribed thickening, diaphragm involvement, CPA involvement, and the extent of the thickening (Broderick et al., 1992). Multivariate linear regression, allowing for control of potential confounders, found decreased FVC was significantly related to DPT, plaques, CPA involvement, and extent of the thickening, but not diaphragmatic involvement (Broderick et al., 1992).

The mechanisms for reduced lung volume in individuals with asbestos-related DPT have been examined by measuring lung function and changes in diaphragm length, rib-cage dimensions, and subphrenic volume in 26 patients during breathing (Singh et al., 1999). DPT reduced both total lung capacity and FVC with corresponding decreases in rib-cage expansion and movement of the diaphragm, consistent with the restrictive nature of these lesions, which may encase part of the lung (Singh et al., 1999). These direct measurements of the effect of DPT chest wall and diaphragmatic motion illustrate the role of DPT in reducing lung volume, contributing to restrictive deficits in pulmonary function. Taken together, the epidemiologic evidence and the mechanistic information that support a restrictive effect of fibrotic lesion in the visceral pleura, substantiate the associations between DPT and decreased pulmonary function. As such, the observation of DPT on standard radiographs is representative of pathological changes directly related to reduced lung function and is, therefore, an indication of adversity, and, can serve as an appropriate health endpoint for consideration in RfC derivation.

5.2.2.3. *Health Effects of Localized Pleural Thickening (LPT) Viewed on Standard Radiographs*

Localized pleural thickening (LPT) viewed on a standard radiograph may include both pleural plaques and pleural thickening that does not involve blunting of the costophrenic angle (ILO, 2002). Thus, both parietal plaques and localized thickening of the visceral pleura may be designated as LPT. Thickening of the parietal pleura is due to an acellular collagen plaque (basket weave of collagen fibers) between the parietal pleura and the ribcage (or along the diaphragm) often described as discrete or circumscribed pleural plaques (ATS, 2004; Jones,

2002). Thickening of the visceral pleural is a fibrosis with diffuse borders and may extend into the lung parenchyma (ATS, 2004; Jones, 2002). The pathology and health effects of the different lesions are evaluated here in the characterization of the health significance of LPT.

Costal parietal plaques occur between the thoracic cage and parietal pleura, which is normally adherent to the thoracic cage (ATS, 2004; Jones, 2002). Costal parietal plaques have been described as collagen deposits with ragged irregular edges and up to 1 cm in depth and may be calcified. These parietal plaques have been associated with constricting pain in the thoracic cavity (Mukherjee et al., 2000). The parietal pleura is well innervated by the intercostal and phrenic nerves and is considered very sensitive to painful stimuli (Jones, 2002). With respect to parietal plaques, pain during exertion or exercise could result in restrained chest wall motion during exertion or exercise. Thus, Bourbeau et al. (1990) hypothesized that the dyspnea and changes in pulmonary function noted in individuals with pleural plaques may be due to physical irritation and perhaps a constricting action where parietal plaques are well progressed or numerous and impact a large proportion of the parietal surface.

Kouris et al. (1991) examined the presence of dyspnea, and measures of pulmonary function (i.e., FVC, FEV1, and FEV%[27]) in asbestos-exposed workers ($n = 913$) in relation to radiographic signs of lung and pleural anomalies. Radiographs were contemporary to the study and read in accordance with ILO (1980) guidelines. Pleural plaques were associated with reduced FVC and FEV1.0 (87.6% and 84.1% of predicted, respectively, $p < 0.0005$), although deficits associated with diffuse thickening were greater (76.4% and 73.9%, $p < 0.0005$) (Kouris et al., 1991). Correspondingly odds ratios for decreased FVC and FEV1.0 (80% decrement) were increased by the presence of both plaques and diffuse thickening (1.5 for plaques and 4.2 and 4.7 for diffuse thickening, respectively). Interestingly, when history of lung disease was considered, pleural plaques had a greater effect in individuals without previous lung disease (OR of 2.1 for FVC and 1.7 for FEV1.0).

Pleural thickening in general is associated with decreased pulmonary function (Petrovic et al., 2004; Wang et al., 2001; Miller et al., 1994)and this association is strengthened as the severity of the pleural thickening increases (Lilis et al., 1991). Few available studies have examined the relationship between pleural plaques identified on standard radiographs (ILO,

[27]Forced Vital Capacity (FVC); Forced Expiratory Volume in 1 second (FEV1) and Percent FVC (FEV%) = [(100 × FEV1) ÷ FVC].

1980) and pulmonary function without including DPT in the analysis and adequately controlling for the presence of small opacities (indicative of parenchymal damage)[28].

Lilis et al. (1991) examined pulmonary function in long-term asbestos insulation workers, and found that one measure (FVC) decreased significantly as the severity of pleural fibrosis (all types, as indicated by a pleural index) increased. This decrease was more dramatic when including parenchymal changes (small opacities) or if DPT was viewed separately. A second analysis focusing on participants with pleural plaques found an inverse relationship between severity of the pleural plaques and FVC ($p < 0.0001$), when adjusting for the independent effects of duration, smoking and presence of small opacities (Lilis et al., 1991). This finding supports a view that pleural plaques, when extensive, may contribute to restrictive lung deficits, but the analysis included individuals with known small opacities (e.g., lung fibrosis). The authors do not address the potential that the pleural index may also correspond to increased severity of parenchymal changes, potentially confounding the analysis where accounting for small opacities (profusion scores of 1/0 or greater) may not adequately control for asbestos-related parenchymal damage.

Oliver et al. (1988) studied the relationship between pulmonary function and pleural plaques in asbestos-exposed railway workers ($n = 383$). Case selection included exclusion of workers with DPT (ILO, 1980) and exclusion of any indication of small opacities (only profusion scores of 0/0 were included). Standard spirometry was conducted to evaluate restrictive and obstructive pulmonary deficits. Additionally, single-breath diffusing capacity (DLCO) was measured which would indicate parenchymal defects. The DLCO was similar in subjects with and without circumscribed plaques, suggesting little or no subradiographic parenchymal damage, which corresponded to the presence of pleural plaques. Pleural plaques were associated with both decreased FVC and pulmonary restriction ($p = 0.03$ and 0.04, respectively) where the diagnostic certainty for the plaques was considered 'definite', and there was an association between level of diagnostic certainty and these pulmonary deficits ($p = 0.02$) (Oliver et al., 1988). Quantitative pleural score, based on the number and extent of plaques, was

[28]It is difficult to control for effects subradiographic parenchymal fibrosis on lung function, where it may not have progressed to visible small opacities, and it has been suggested that reduced lung function, which has been associated with circumscribed plaques in some studies, may be reflecting the effects of subradiographic parenchymal changes, rather than a direct effect of DPP (ATS, 2004; Erdinç et al., 2003; Miller and Zurlo, 1996; Broderick et al., 1992).

also associated with decreased FVC and pulmonary restriction ($p = 0.0135$ and 0.0126, respectively) (Oliver et al., 1988). Of the available studies that assess pleural thickening with standard radiographs, this study best controls for the possibility of subradiographic parenchymal damage and is, therefore, strong evidence that circumscribed pleural plaques independently impact pulmonary function. The observed restrictive pulmonary deficit is consistent with the potential for pleural plaques to restrict chest wall motion or the elasticity of the diaphragm.

Three high-resolution computed tomography (HRCT) studies were conducted specifically to assess the potential for parietal plaques to impact lung function. Staples et al. (1989) report no difference in lung function or diffusing capacity between participants ($n = 76$) with and without pleural plaques. Soulat et al. (1999) found no difference in FEV1 or FVC between asbestos-exposed insulators with ($n = 84$) and without ($n = 51$) pleural plaques in the absence of any parenchymal changes. As severity of pleural thickening has been shown to be positively associated with decrease measures of pulmonary function, Van Cleemput et al. (2001) not only examined the effect of HRCT defined pleural plaques on pulmonary function, but also assessed the extent of the pleural plaques. Neither the presence nor extent of pleural plaques were associated with lung function parameters (diffusing capacity or normalized spirometric values) (van Cleemput et al., 2001). Where pleural plaques and diffuse thickening (visceral pleura) were both identified by HRCT and correlated to pulmonary function, diffuse visceral thickening—but not plaques—were associated with decreased lung volume and FVC (Copley et al., 2001). Although CPA involvement was not independently assessed, several scoring systems for severity were compared which included CPA involvement, and as in other studies, increased severity correlated to greater decrements.

The mechanisms for reduced lung volume in individuals with asbestos-related pleural plaques and DPT have been examined by measuring lung function and changes in diaphragm length, rib-cage dimensions and subphrenic volume in 26 patients during breathing (Singh et al., 1999). Pleural plaques alone did not reduce any of the measures of lung function in this study, but there were indications of reduced diaphragm movement (Singh et al., 1999). This may be an indication that diaphragmatic plaques in the parietal pleura have the potential to attenuate the movement of the diaphragm during breathing. Because this study is relatively small ($N = 26$) and a distinction was not made between costal and diaphragmatic plaques by the study authors,

additional work is needed to better understand the direct effects of pleural plaques on lung function.

Although some researchers have questioned that pleural plaques alone directly impact pulmonary function, a critical review of the literature from 1965-1999 concludes: "1) Individuals with asbestos-induced pleural plaques may have alterations in pulmonary function and /or clinical symptoms that are independent of smoking and radiographic parenchymal fibrosis and, 2) the respiratory changes dues to asbestos-induced pleural plaques are generally less severe than those caused by pleural thickening" (Rockoff et al., 2002). Therefore, although the evidence is mixed, pleural plaques may be independently associated with reduced pulmonary function.

No studies correlating pulmonary function to radiographic signs of localized pleural thickening (LPT) using the ILO (ILO, 2002) guidelines could be located. However, several researchers employed similar classification schemes, modifying earlier ILO classification systems, such that DPT was diagnosed only in conjunction with blunting of the CPA. This modification potentially includes cases of diffuse pleural thickening (without CPA blunting) in their analysis of pleural plaques, making their findings somewhat applicable to the current classification of LPT (García-Closas and Christiani, 1995; Broderick et al., 1992). Pleural thickening (without CPA blunting) was associated with mixed respiratory impairment in a study of asbestos-exposed construction carpenters (n = 631) (OR of 3.7 [95% Confidence Interval (CI): 1.4–12.3]) but was only weakly associated when the outcome was restrictive deficit specifically (1.3 [95% CI: 0.4–3.9]) (García-Closas and Christiani, 1995). Broderick et al. (1992) found decreased FVC was not only significantly associated with "diffuse thickening" (with CPA blunting) but also with "pleural plaques" (which included all pleural thickening without CPA blunting). The severity of pleural thickening (both as width or percentage of lateral wall) and calcification was associated with reduced FVC as well (Broderick et al., 1992). Kilburn and Warshaw (1991) assessed pulmonary function in individuals with "plaques only," "diffuse thickening only," and "diffuse thickening with CPA blunting," showing progressive deficits across these categories in FVC, FEV1, and mid-expiratory flow (e.g., FEV1: 90.5, 86.2, and 49.4% [$p < 0.05$], respectively). Again, there is a trend that diffuse thickening has a greater impact on lung function parameters, although an independent effect of plaques cannot be ruled out by these data.

In summary, the radiographic classification of localized pleural thickening (LPT) under current ILO guidelines may include both parietal plaques (in the pleura lining the interior of the ribcage) and diffuse visceral thickening (without CPA obliteration) (ILO, 2002). The two lesions (parietal plaques and localized visceral thickening) are distinct and may contribute independently to observed health effects. Parietal plaques are known to induce chronic constricting chest pain that increases in severity as the extent of the plaques increases. Pleural thickening in general is associated with reduced lung function parameters with increased effect correlating with increased severity of the pleural thickening (Petrovic et al., 2004; Wang et al., 2001; Miller et al., 1994; Lilis et al., 1991). There is clear evidence from HRCT studies that the presence and extent of visceral thickening does impair lung function, although, when evaluated independently, parietal plaques were not statistically correlated with decreased pulmonary function (Copley et al., 2001; Schwartz et al., 1993). Specifically considering the designation of LPT, lung function impairment has been demonstrated in several studies where pleural thickening without CPA involvement has been studied (García-Closas and Christiani, 1995; Broderick et al., 1992; Kilburn and Warshaw, 1991). Thus, the radiographic classification of localized pleural thickening (LPT) (ILO, 2002) includes pleural lesions associated with chronic chest pain, decreased lung volume, and decreased measures of lung function. Therefore, EPA considers LPT an adverse effect and an appropriate endpoint for RfC derivation.

5.2.3. Methods of Analysis

5.2.3.1. *Exposure Data and Choice of Exposure Metric*

EPA collaborated with a research team at the University of Cincinnati to update the exposure reconstruction for use in the job-exposure matrix (JEM) for all workers in the Marysville, OH cohort, taking into account additional industrial hygiene data that were not available for previous studies conducted in this cohort. As discussed in detail in Appendix F, exposure estimates for each worker in the O.M. Scott Marysville, OH plant were developed based on available industrial hygiene data from the plant. Figure 5-1 shows the average exposure concentrations of fibers in air (PCM fibers/cc)[29] of each department from 1957 to 2000,

[29]PCM, where fibers are viewed and counted by light microscopy, does not identify the composition of the fiber. Thus, the mineralogy of fibers identified under PCM cannot be determined.

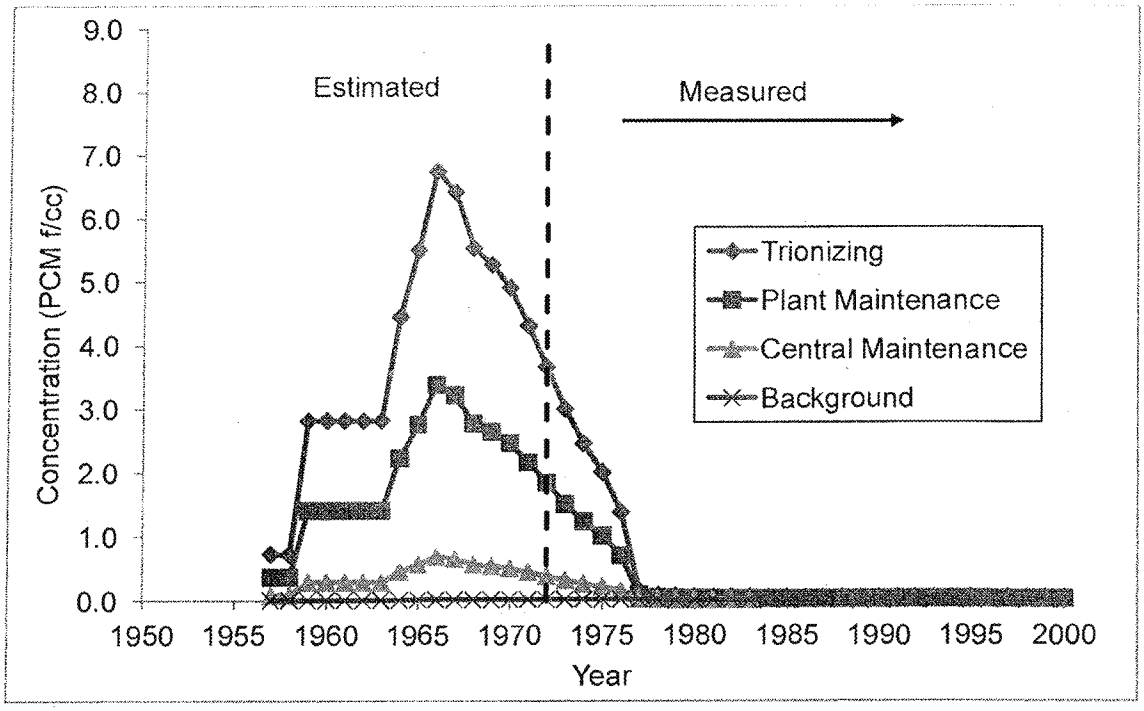


Figure 5-1. Estimated and measured exposure concentrations in Marysville, OH facility[a]

[a]Trionizing is a term used in the Marysville, OH facility and includes unloading of rail cars containing vermiculite ore (track), using conveyers to move the vermiculite ore into the expander furnaces, separation of the expanded vermiculite from sand, blending in of lawn care chemicals, and drying and packaging of the final product. As no unexpanded ore was used in pilot plant, research, polyform, office, packaging, or warehouse, jobs in these categories were assigned as background. Workers assigned to plant maintenance activities spent 50% of their time in trionizing areas and 50% of their time in areas assigned as plant background. Workers assigned to central maintenance spend 10% of their time in trionizing areas and 90% of their time in areas assigned as plant background. Central maintenance jobs were eliminated in 1982 and contracted out (see Appendix F).

indicating the time periods when fiber measurements were not available ('Estimated') and were available ('Measured').

In brief, the starting point for the JEM was the measured or estimated concentration of fibers in air (fibers/cc) of each department from 1957–2000. The distribution of exposure by department is summarized in Figure 5-1. Using available data on the year of hire and the departments in which each person worked, the cumulative exposure (fibers/cc-year) for each worker for each year since the date of hire was estimated. Each worker's cumulative exposure was then adjusted to a cumulative human equivalent exposure for continuous exposure (CHEEC; fibers/cc-year) to represent exposure 24 hours/day and 365 days/year (assuming that any exposure off site was zero) for the full duration of employment. Adjustments for different inhalation rates in working versus nonworking time periods were incorporated in this analysis. The calculated value is similar to what EPA usually refers to as continuous human equivalent exposure (U.S. EPA, 1994b). These calculations are somewhat more complex than the usual conversions to equivalent continuous exposure concentrations that EPA makes in the analysis of occupational studies. Conversions for noncancer effects are usually made using an adjustment factor of 240 days ÷ 365 days × 10 m^3 ÷ 20 m^3 (U.S. EPA, 1994b). However, the adjustment factor in this current assessment takes into account the extensive seasonal overtime for some job codes at the Marysville facility, as well as other annual periods when work hours were reduced (see Appendix F). The estimated CHEEC was used to represent Libby Amphibole asbestos exposure in all subsequent analyses because it combines aspects of both intensity of exposure and duration of exposure.[30] For Libby Amphibole asbestos, the exposure metric is calculated as cumulative exposure (fibers/cc-year). Cumulative exposure is a commonly evaluated exposure metric in occupational studies, especially for mineral fibers, where fiber retention may be relevant to toxicity. It should be noted that discrete parietal plaques have often been associated with other exposure metrics (e.g., mean exposure, TSFE) (i.e., Paris et al., 2008; Jakobsson et al., 1995; Ehrlich et al., 1992; Copes et al., 1985). Paris et al. (2008) show significant exposure-response relationships for both mean and cumulative exposure metrics for pleural plaques (identified by HRCT) among workers with mixed fiber exposures, when accounting for age, smoking, and TSFE. Mean exposure provided a better overall fit (Paris et al., 2009). Thus, EPA has conducted an uncertainty assessment for the RfC derivation from the sub-cohort by also

[30]The University of Cincinnati used the term CHEEC in its report (see Appendix F).

exploring alternative methods to weight the $BMCL_{10}$ in units of cumulative exposure, to represent the average exposure needed for RfC derivation (see Section 5.3.7).

Because localized pleural thickening does not generally occur immediately after exposure and requires some time to develop to the state that it can be detected on a conventional chest X-ray, exposures that occur close to the time of X-ray may not contribute to the occurrence of observable disease and may obscure the exposure-response relationship. Accordingly, a lagged exposure (i.e., cumulative exposure discounting the most recent time period) may be the most appropriate measure to use. Therefore, exposure estimates with various lags were investigated (lags of 0, 5, 10, 15, and 20 years). For example, a CHEEC value based on a lag of 5 years excludes all exposures that occurred within 5 years of the date of X-ray. Looking at the occurrence of the outcome for various categories of time elapsed since first exposure, the first localized pleural thickening was detected ~10 years after the first exposure.

5.2.3.2. *Data Sets for Modeling Analyses*

The individual health outcome data for all workers who participated in the Lockey et al. (1984) study and the follow-up study by Rohs et al. (2008) were used for exposure-response modeling. To avoid any bias from previous occupational exposure to asbestos, only the data from those who did not report any previous occupational exposure to asbestos were used. The data from Lockey et al. (1984) and Rohs et al. (2008) were combined for the full cohort to provide a greater range in time from first exposure (described below). Outcome assessments, i.e., chest X-rays, were performed at two different time points, 1980 and 2002–2005. While the evaluation approaches were generally similar (independent readings by three certified B-readers), it is important to note that X-ray readings were performed by different individuals, under a different reading protocol in 1980 (modified 1971 ILO standards) compared to 2000s [ILO (2002) standards], leading to some uncertainty in statistical analyses that combine these data sets. An additional consideration is human body composition—in some cases, difficulty in distinguishing fat pads from true pleural thickening may lead to misclassification of the outcome. BMI measurements are available for the latter study but not for the 1980 evaluation; the effect of BMI was investigated and is discussed below.

Radiographs were evaluated by two B-readers with a consensus evaluation by a third reader in the case of disagreement in the original study by Lockey et al. (1984). In the follow-up

by Rohs et al. (2008), a radiographic reading was considered positive "when the median classification from the three independent B readings was consistent with pleural and/or interstitial changes" (p. 631). Because the ILO criteria were updated in 2000, the reader forms from Lockey et al. (1984) showing pleural changes were evaluated for consistency with the ILO 2000 criteria. This reevaluation did not result in any change in the diagnosis for any individual from the 1980 reading.[31] In addition, no difference in reported X-ray quality was noted between the Lockey et al. (1984) data and the follow-up by Rohs et al. (2008).

The full data set of the exposure-response relationship for localized pleural thickening was as follows. The radiographic data from Lockey et al. (1984) (n = 513) and Rohs et al. (2008) (n = 280), were combined for a total of 793 X-ray evaluations (this includes repeated X-rays on the same individual). X-rays obtained from workers who reported exposure to asbestos at other locations were excluded from consideration (n = 793 – 105 = 688 X-ray evaluations).

For workers who were X-rayed in both Lockey et al. (1984) and Rohs et al. (2008), one of the observations was excluded so that there were no repeat observations for individual workers in the data set used for modeling. For workers who were negative for localized pleural thickening in Lockey et al., the (1984) study data were excluded, and the Rohs et al. (2008) data were retained. For workers who were positive for localized pleural thickening in Lockey et al. (1984) and also in Rohs et al. (2008), the 1984 study data were retained. One worker was positive in 1984 and negative in 2008 (removing this worker from the analysis did not change results). The 2008 study data were retained for this worker. This procedure resulted in n = 688 X-rays – 252 duplicates = 436 X-rays, representing 436 individual workers.

Two workers from Lockey et al. (1984) were excluded because the start day and the X-ray date were the same (n = 436 – 2 = 434). For each worker, the estimated cumulative exposure corresponded to the date of the X-ray retained for analysis—if the 1980 X-ray was used, the individual's cumulative exposure estimate covered the period from start of work through the X-ray date in 1980. If the 2002–2005 X-ray was used, cumulative exposure covered the period from start of work through the date of job stop or 2000, whichever occurred earlier.

[31]Personal communication (e-mail) from Dr. James Lockey, University of Cincinnati, to Dr. Robert Benson in March 2011 reports that a review of the 1980 B-reader forms using the ILO 2000 guidelines would not result in changes in individual diagnosis for study participants.

The Marysville cohort data comprise 434 workers who were not previously exposed to asbestos and had at least one X-ray observation. Because the concentration of Libby Amphibole asbestos in workplace air was estimated rather than measured for all years prior to 1972, this data set was stratified into two subsets: (1) workers hired in 1972 or after (for whom all exposure values are measured), and (2) workers hired before 1972 (for whom some of the exposure values are estimated). Distributions of cases and TSFE (T) at each outcome assessment are shown in Table 5-3.

Table 5-3. Distribution of cases and time from first exposure (T) for cohort of Marysville workers

	All participants[a]		First exposed before 1972		First exposed 1972 or later	
	Cases/Total	Range of T	Cases/Total	Range of T	Cases/Total	Range of T
Examined 1980 (Lockey et al., 1984)	5/434	0.42–23.43	4/236	8.75–23.43	1/198	0.42–8.42
Examined 2002–2005 (Rohs et al., 2008)	57/252	23.14–47.34	45/133	31.07–47.34	12/119	23.14–32.63
Marysville cohort (n = 434, examination in either 1980 or 2002–2005)	61/434	0.42–47.34	48/236	8.75–47.34	13/198	0.42–32.63

[a]The 252 individuals examined in 2002–2005 were also examined in 1980. Note that there were originally 513 individuals in the Lockey et al. (1984) cohort; of these, 77 had previous asbestos exposure and were excluded (n = 436). Two individuals were excluded because their X-ray date was the same as their employment start date (n = 434). These exclusions are also reflected in the Rohs et al. (2008) cohort.

Source: Rohs et al. (2008) and Lockey et al. (1984).

The more accurate exposure data are considered to be those from 1972 and later, as these data were based on analytical measurements. Due to the longer follow-up time and additional covariate information, the most informative outcome data come from the 2002–2005 examination. Based on these considerations, a sub-cohort of the Marysville workers, which includes data from workers in the 2002–2005 examination, and who began work in 1972 or later

(12 cases of localized pleural thickening and 106 unaffected individuals[32]) (Rohs et al., 2008), was chosen as the preferred analysis to develop a point of departure (POD) for localized pleural thickening to serve as the basis for the RfC. Additionally, sample POD estimates based on statistical analyses of results from the full cohort [Lockey et al. (1984) and Rohs et al. (2008) combined, as described above] were included for comparison.

5.2.3.3. *Statistical Modeling of the Sub-cohort*

EPA performed analyses of study results for the sub-cohort whose exposures began on or after 1/1/1972 when workplace PCM measurements were available, reducing uncertainties associated with exposure assessment. Localized pleural thickening (LPT), as diagnosed from a standard radiograph (ILO, 2002), was selected as the critical effect based on the health effects associated with pleural thickening specific to this diagnosis (see Section 5.2.2.3). Alternative critical effects were not considered for the sub-cohort analysis given the limited number of cases (one case of DPT and no cases of small opacities). Epidemiologic methods were used to analyze the exposure-response data, and benchmark concentration (BMC) methodology was used to estimate PODs. In this approach, the available data are fit to a set of mathematical exposure-response models to determine an appropriate empirical representation of the data. General model fit is evaluated to determine whether the model form appropriately represents the data; here, this was done using the Hosmer-Lemeshow test (a form of the Pearson χ^2 goodness-of-fit statistic). Among models with adequate general fit, a recommended model form is then determined; commonly, this is the model with the best fit as measured by Akaike's Information Criterion (AIC) value among these model forms judged to provide an appropriate and statistically adequate representation of the data. For inhalation data, the BMC is defined as the exposure level, calculated from the best-fit model, which results in a specified benchmark response (BMR). The RfC is derived from the lower 95% confidence limit of the BMC, referred to as the BMCL, which accounts for statistical uncertainty in the model fit to the data. All

[32]There was one individual whose radiographic examination indicated diffuse pleural thickening, who was excluded from further analyses of the preferred sub-cohort. Diffuse pleural thickening represents a more severe outcome than the selected critical effect of LPT—including this individual as a case would not be appropriate given that the critical effect is selected to represent a most sensitive endpoint, and the subsequent selection of a benchmark response in modeling efforts. Diffuse pleural thickening is considered separately as an endpoint (with appropriate benchmark response) in sensitivity analyses of alternative outcomes in the larger group of workers examined in 2002–2005 (see Section 5.3.8).

analyses were performed using SAS® statistical software v. 9.1. BMCLs were obtained by the profile likelihood method as recommended by Crump and Howe (1985) using the NLMIXED (nonlinear mixed modeling) procedure in SAS (Wheeler, 2005) (see Appendix E for details).

For models where a background parameter is included, a 1% risk of localized pleural thickening was assumed. Establishing a background rate for LPT prevalence is problematic for several reasons. Little data exist to define background rates for LPT, as this designation is more recent, and the majority of the published data use earlier ILO guidelines, which define discrete pleural plaques (DPP). Secondly, it is difficult to define a population without exposure to asbestos in any setting. As environmental and community exposures can increase pleural thickening (Weill et al., 2011; Luo et al., 2003; Hiraoka et al., 1998; Zitting et al., 1996) the question arises, Is there a true background rate? Also, in general, pleural thickening increases with both age and TSFE in a population. There is a study that reports the LPT in Libby community members with no reported pathways of exposure (Weill et al., 2011). LPT prevalence is reported at 0.4% in participants age 25–40, and 1.4% in participants age 41–50 (based on X-rays taken in 2000). Older study participants (61–90) had a LPT prevalence of 12.7%, likely influenced by high historical exposures, as well as the increased TSFE. In two studies of persons not known to be previously exposed to asbestos, Anderson et al. (1979) and Castellan et al. (1985) report DPP estimated prevalence of 1.2% (4/326) and 0.2% (3/1,422), respectively. In cross-sectional studies, which may include persons with occupational exposure to asbestos, Rogan reported DPP prevalence estimates of 1.2% in the National Health and Nutrition Examination (NHANES) I study (1971–1975) (Rogan et al., 1987) and 3.9% in the NHANES II study (Rogan et al., 2000). Among military populations, two studies have reported an estimated DPP prevalence of 2.3% (Muller et al., 2005; Miller and Zurlo, 1996). Based on these reports, the 1% background rate was chosen as representing the prevalence among persons without occupational exposure to asbestos in the age range of the Rohs et al. (2008) study population. As there is some uncertainty regarding the true background rate for LPT, a sensitivity analysis was performed where the model includes the background rate as an estimated parameter rather than using the set value of 1%. There was little change in the resulting model fits or BMCLs (see Section 5.3.4).

In the absence of agent-specific information to assist in identifying a BMR, a 10% extra risk was judged to be a minimally biologically significant level of change, and is also

recommended for standard reporting purposes (U.S. EPA, 2000a). LPT is an irreversible pathological change and associated with health effects including chronic pain, dyspnea, and deficits in pulmonary function (see Section 5.2.2.3). The likelihood and severity of these health effects increases with increased extent and severity of the pleural thickening. However, as the data from the critical study do not provide information on the severity of the lesions, we cannot assess the relative likelihood of any of these health effects. Thus, the observed LPT prevalence may include a range of lesions from minimally adverse to severe. The biology of more severe lesions (i.e., DPT and small opacities) could justify lower BMRs; however, there are not enough cases to model these endpoints in this sub-cohort. A sensitivity analysis was conducted using the data set included in Rohs et al. (2008) to examine the impact of choice of BMR and critical effect on the POD (see Section 5.3.8).

5.2.3.3.1. *Statistical model evaluation and selection*

Dichotomous statistical models describing the probability of individual response as a function of cumulative exposure (represented by CHEEC in units of fibers/cc-year) were used. In order to investigate the key explanatory variables for analysis, a forward-selection process was used to evaluate the association of each of the potential covariates with the risk of localized pleural thickening, controlling for Libby Amphibole asbestos exposure. Covariates considered for inclusion in the model were TSFE (T), age at X-ray, gender, smoking history, and BMI. This initial modeling was done using a standard logistic regression model, as is commonly applied in analysis of epidemiological data. The base model was a logistic regression model with cumulative Libby Amphibole asbestos exposure (natural log transformed) as the independent variable. This model provided an adequate fit to the data (Hosmer-Lemeshow p-value of 0.64), and the exposure variable was statistically significantly associated with the outcome (beta = 0.5676, standard error, [SE] = 0.2420 increase in log odds for every unit increase in CHEEC, p-value = 0.02). Covariates were evaluated according to whether inclusion of the covariate improved model fit as assessed by the AIC, and statistical significance of the covariate. When controlling for Libby Amphibole asbestos exposure, none of these covariates were associated with odds of localized pleural thickening: T: p-value = 0.89; age at X-ray: p-value = 0.77; gender: p-value = 0.78; smoking history: p-value = 0.17; BMI: p-value = 0.41. The inclusion of each of the covariates with the exception of smoking increased the AIC for the

APPENDIX C – 7

http://www.epa.gov/risk/health-risk.htm

EPA United States Environmental Protection Agency

Risk Assessment

Human Health Risk Assessment

Introduction

A human health risk assessment is the process to estimate the nature and probability of adverse health effects in humans who may be exposed to chemicals in contaminated environmental media, now or in the future.

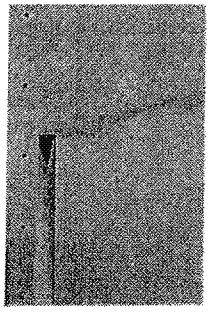
To explain this better, a human health risk assessment addresses questions such as:

- What types of health problems may be caused by environmental stressors such as chemicals and radiation?
- What is the chance that people will experience health problems when exposed to different levels of environmental stressors?
- Is there a level below which some chemicals don't pose a human health risk?
- What environmental stressors are people exposed to and at what levels and for how long?
- Are some people more likely to be susceptible to environmental stressors because of factors such as age, genetics, pre-existing health conditions, ethnic practices, gender, etc.?
- Are some people more likely to be exposed to environmental stressors because of factors such as where they work, where they play, what they like to eat, etc.?

The answers to these types of questions helps decision makers, whether they are parents or public officials, understand the possible human health risks from environmental media.

How does EPA conduct a Human Health Risk Assessment?

Human health risk assessment includes 4 basic steps, and is generally conducted following various EPA guidance documents.



Planning - Planning and Scoping process
EPA begins the process of a human health risk assessment with planning and research.

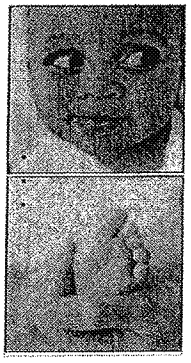
Step 1 - Hazard Identification
Examines whether a stressor has the potential to cause harm to humans and/or ecological systems, and if so, under what circumstances.

Step 2 - Dose-Response Assessment
Examines the numerical relationship between exposure and effects.

Step 3 - Exposure Assessment
Examines what is known about the frequency, timing, and levels of contact with a stressor.

Step 4 - Risk Characterization
Examines how well the data support conclusions about the nature and extent of the risk from exposure to environmental stressors.

Why does EPA evaluate whether children may be at greater health risks than adults?



Almost 500 years ago Paracelsus (1493-1541) wrote: "Dosis facit venenum" or "the dose makes the poison." The relationship between dose and response (health effect) is still one of the most fundamental concepts of toxicology - or is it? For pollutants that act as developmental toxicants, the same dose that may pose little or no risk to an adult can cause drastic effects in a developing fetus or a child. Methyl mercury is but one example of a chemical that is much more toxic early in life. Scientists have become increasingly aware that children may be more vulnerable to environmental exposures than adults because:

their bodily systems are developing;
they eat more, drink more, and breathe more in proportion to their body size; and
their behavior, such as crawling and hand-to-mouth activity, can expose them more to chemicals and microorganisms.

In light of what is now known about the greater susceptibility early in life to some stressors, Executive Order 13045 -- Protection of Children from Environmental Health Risks and Safety Risks -- was issued in 1997. This Executive Order directs that all federal agencies, including EPA, shall make it a high priority to identify and assess environmental health risks and safety risks that may disproportionately affect children; and shall ensure that their policies, programs, activities, and standards address disproportionate risks to children that result from environmental health risks or safety risks.

Note: To assist scientists in assessing risks specifically to children, EPA has developed A Framework for Assessing Health Risk of Environmental Exposures to Children along with specific guidance to risk assessors including Guidance on Selecting Age Groups for Monitoring and Assessing Child-Hood Exposures to Environmental Contaminants and Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens.

Last updated on Tuesday, July 31, 2012

EPA United States Environmental Protection Agency

Risk Assessment

Basic Information

Before finding out about risk assessment there are some fundamental principles you need to understand:

What is risk? What is a stressor?

While there are many definitions of the word risk, **EPA considers risk to be the chance of harmful effects to human health or to ecological systems** resulting from exposure to an environmental stressor.

A stressor is any physical, chemical, or biological entity that can induce an adverse response. Stressors may adversely affect specific natural resources or entire ecosystems, including plants and animals, as well as the environment with which they interact.

What is risk assessment?

EPA uses risk assessment to characterize the nature and magnitude of health risks to humans (e.g., residents, workers, recreational visitors) and ecological receptors (e.g., birds, fish, wildlife) from chemical contaminants and other stressors, that may be present in the environment. Risk managers use this information to help them decide how to protect humans and the environment from stressors or contaminants. Note that "risk managers" can be:

- federal or state officials whose job it is to protect the environment,
- business leaders who work at companies that can impact the environment, or
- private citizens who are making decisions regarding risk.

At EPA, environmental risk assessments typically fall into one of two areas:

- Human Health
- Ecological

Risk assessment is, to the highest extent possible, a scientific process. In general terms, risk depends on the following factors:

- How much of a chemical is present in an environmental medium (e.g., soil, water, air),
- How much contact (exposure) a person or ecological receptor has with the contaminated environmental medium, and
- The inherent toxicity of the chemical.

Following a planning and scoping stage where the purpose and scope of a risk assessment is decided, the risk assessment process usually begins by collecting measurements that characterize the nature and extent of chemical contamination in the environment, as well as information needed to predict how the contaminants behave in the future. Here are some useful links to get started:

- EPA's Guidance on Planning and Scoping
- Planning a human health risk assessment
- Planning an ecological risk assessment

Based on this, the risk assessor evaluates the frequency and magnitude of human and ecological exposures that may occur as a consequence of contact with the contaminated medium, both now and in the future.

This evaluation of exposure is then combined with information on the inherent toxicity of the chemical (that is, the expected response to a given level of exposure) to predict the probability, nature, and magnitude of the adverse health effects that may occur. In the ideal world, all risk assessments would be based on a very strong knowledge base (i.e., reliable and complete data on the nature and extent of contamination, fate and transport processes, the magnitude and frequency of human and ecological exposure, and the inherent toxicity of all of the chemicals). However, in real life, information is usually limited on one or more of these key data needed for risk assessment calculations. This means that risk assessors often have to make estimates and use judgment when performing risk calculations, and consequently all risk estimates are uncertain to some degree. For this reason, a key part of all good risk assessments is a fair and open presentation of the uncertainties in the calculations and a characterization of how reliable (or how unreliable) the resulting risk estimates really are.

Developing a risk assessment is often an iterative process, which involves researchers identifying and filling data gaps in order to develop a more refined assessment of the risk. This in turn may influence the need for risk assessors and risk managers to refine the scope of the risk assessment further triggering the need for more data or new assumptions.

What is risk management?

As described in EPA's Risk Characterization Handbook (PDF) (89 pp, 8.9MB, about PDF), "Risk Management" is the process which evaluates how to protect public health. Examples of risk management actions include deciding how much of a substance a company may discharge into a river; deciding which substances may be stored at a hazardous waste disposal facility; deciding to what extent a hazardous waste site must be cleaned up; setting permit levels for discharge, storage, or transport; establishing national ambient air quality standards; and determining allowable levels of contamination in drinking water.

Risk assessment provides "INFORMATION" on potential health or ecological risks, and risk management is the "ACTION" taken based on consideration of that and other information, as follows:

- Scientific factors provide the basis for the risk assessment, including information drawn from toxicology, chemistry, epidemiology, ecology, and statistics - to name a few.
- Economic factors inform the manager on the cost of risks and the benefits of reducing them, the costs of risk mitigation or remediation options and the distributional effects.
- Laws and legal decisions are factors that define the basis for the Agency's risk assessments, management decisions, and, in some instances, the schedule, level or methods for risk reduction.
- Social factors, such as income level, ethnic background, community values, land use, zoning, availability of health care, life style, and psychological condition of the affected populations, may affect the susceptibility of an individual or a definable group to risks from a particular stressor.
- Technological factors include the feasibility, impacts, and range of risk management options.
- Political factors are based on the interactions among branches of the Federal government, with other Federal, state, and local government entities, and even with foreign governments; these may range from practices defined by Agency policy and political administrations through inquiries from members of Congress, special interest groups, or concerned citizens.
- Public values reflect the broad attitudes of society about environmental risks and risk management.

Who evaluates the risks?

The table below outlines which EPA office or other federal agency is responsible for assessing and managing risks associated with particular stressors.

Stressor	EPA Office	Other Federal Agencies
Air Pollution	Office of Air and Radiation	
Hazardous substances, pollutants, and waste	Office of Solid Waste and Emergency Response	
Pharmaceuticals		FDA's Center for Drug Evaluation and Research
Pesticides	Office of Pesticide Programs	U.S. Consumer Product Safety Commission (toys and other consumer products) FDA's Center for Food Safety and Applied Nutrition
Radiation including radon	Radiation Programs	
Toxic substances, human exposure, environmental exposure	Office of Pollution Prevention and Toxics Office of Research and Development	
Vaccines		FDA's Center for Biologics Evaluation and Research
Water pollution	Office of Water	

How does EPA conduct risk assessments?

At EPA, environmental risk assessments typically fall into one of two areas: human health risk assessments or ecological risk assessments. These are described in steps or parts due to the differences in how each of these are conducted at EPA.

Where do I find EPA risk assessments?

Because risk assessments are performed all over EPA (see the EPA Organization Chart for other EPA Offices and Regions), risk assessments are produced by many of EPA's Regions and Program Offices. Here is a list of primary risk assessment sources:

- Integrated Risk Information System (IRIS) Chemical Summaries and Toxicological Reviews
  - What is IRIS?
  - What is the the IRIS Process for chemical assessment?
- National Center for Environmental Assessment (NCEA) Published Assessments
  - Agent-based risk assessments
    - Carbon Monoxide
    - Diesel Exhaust
    - Dioxin
    - Drinking Water and Disinfection By-Products
    - Lead
    - Mercury
    - Nitrogen Oxide (NOx)
    - Ozone
    - Particulate Matter
    - Pesticide Ecological Risk Assessments
    - PCBs
    - Radon in Homes
    - Secondhand Smoke (ETS)

- Sulfur Oxide
- Place-based risk assessments
  - Biological Assessments (Water)
  - National (Water) Assessment Database
  - Watershed and other place based risk assessments

See Tools & Guidance for a list of more resources.

Where can I find additional information on risk assessment for the public?

EPA has posted a few citizen guides that may be of help for those new to risk assessment. Here is a list of available publications:

- U.S. EPA. A Citizen's Guide to Radon: The Guide to Protecting Yourself and Your Family from Radon. EPA 402-K-07-009. May 2007.
- U.S. EPA. Air Pollution and Health Risk. EPA 450/3-90-022. March 1991
- U.S. EPA. Evaluating Exposures to Toxic Air Pollutants: A Citizen's Guide. EPA 450/3-90-023. March 1991.
- U.S. EPA. RCRA: Reducing Risk from Waste. EPA 530-K-97-004. Sept 1997.
- U.S. EPA. Risk Assessment for Toxic Air Pollutants: A Citizen's Guide . EPA 450/3-90-024. March 1991.

What can I do? Participating in risk assessments

- A Community Guide To Superfund Risk Assessment--What It's All About And How You Can Help
  In Spanish: De qué se trata la evaluación de los riesgos y cómo nos puede ayudar
- Superfund Today: Focus on Revisions to Superfund's Risk Assessment Guidance (1999) (PDF) (2 pp., 50K)
- Regional Vulnerability Assessment (ReVA) Decision Toolkit
- Risk-Screening Environmental Indicators (RSEI) Screening Tool

What does EPA mean by "variability","uncertainty", and "probabilistic modeling"?

Consideration must be given to two important factors throughout the development of a risk assessment: variability and uncertainty.

Variability - Refers to the range of toxic response or exposure. For example, the dose that might cause a toxic response can vary from one person to the next depending on factors such as genetic differences, preexisting medical conditions, etc. Exposure may vary from one person to the next depending on factors such as where one works, time spent indoors or out, where one lives, how much people eat or drink, etc.

Uncertainty - Refers to our inability to know for sure - it is often due to incomplete data. For example, when assessing the potential for risks to people, toxicology studies generally involve dosing of sexually mature test animals such as rats as a surrogate for humans. Since we don't really know how differently humans and rats respond, EPA often employs the use of an uncertainty factor to account for possible differences. Additional consideration may also be made if there is some reason to believe that the very young are more susceptible than adults, or if key toxicology studies are not available. [Learn more about determining uncertainty]

Probabilistic Modeling, a related term, is a technique that utilizes the entire range of input data to develop a probability distribution of exposure or risk rather than a single point value. The input data can be measured values and/or estimated distributions. Values for these input parameters are sampled thousands of times through a modeling or simulation process to develop a distibution of likely exposure or risk. Probabilistic models can be used to evaluate the impact of variability and uncertainty in the various input parameters, such as environmental exposure levels, fate and transport processes, etc.

What is peer review?

Peer review is a documented critical review of a scientific/technical work product which is conducted by scientific experts who are independent of those who performed the work. Peer review can provide an independent evaluation of the assumptions, calculations, extrapolations, alternate interpretations, methodology, acceptance criteria, and conclusions pertaining to the scientific/technical work product.

When evaluating the scientific rigor of our risk assessments, EPA utilizes both standing federal advisory groups of experts such as the Science Advisory Board (SAB) and the FIFRA Scientific Advisory Panel, as well as ad hoc panels to provide peer review. EPA will occasionally seek peer review from outside expert groups such as the National Academy of Science (NAS) for highly complex and/or critical scientific topics.

Last updated on Tuesday, July 31, 2012

APPENDIX C – 8

Tracy B. Horch

From: Schmitt, Addy (USADC) <Addy.Schmitt@usdoj.gov>
Sent: Monday, December 16, 2013 11:06 AM
To: Jayni Lanham
Subject: Beveridge and Diamond v. HHS, 13-1155-JEB
Attachments: B-Reader Form.pdf; CADELAY.doc; CAPILAY.doc; PFTLAY.doc; Questionnaires.doc

Dear Jayni,

I write in response to your email of Thursday evening, December 12, 2013. First, to be clear, my client has already provided all of the information agreed upon by the parties in order to resolve this litigation. We have no obligation to provide additional information, nor do we have any obligation to explain the data you requested and my client provided. Nevertheless, my client is providing the additional information included in and attached to this email as a courtesy - and we trust that you recognize this goes far beyond the terms of the agreement or any obligation to do so. We also trust that you will abide by your agreement to dismiss this case with prejudice by no later than December 20, 2013. Again, my client has gone above and beyond and we do not anticipate any further inquiries or requests before you dismiss the case.

With respect to the occupational categories, my client conducted a search for all instances in which a participant said they did NOT work in a particular job, but for which there were nevertheless start and end dates entered for that job. There were 1,958 records (about 27%) that met this criterion. In other words, that is the data as my client has it.

Regarding the year-of-birth variable, the following code was used:

if 1900<=pbyr<1905 then yrbirth=1;
if 1905<=pbyr<1910 then yrbirth=2;
if 1910<=pbyr<1915 then yrbirth=3;
if 1915<=pbyr<1920 then yrbirth=4;
if 1920<=pbyr<1925 then yrbirth=5;
if 1925<=pbyr<1930 then yrbirth=6;
if 1930<=pbyr<1935 then yrbirth=7;
if 1935<=pbyr<1940 then yrbirth=8;
if 1940<=pbyr<1945 then yrbirth=9;
if 1945<=pbyr<1950 then yrbirth=10;
if 1950<=pbyr<1955 then yrbirth=11;
if 1955<=pbyr<1960 then yrbirth=12;
if 1960<=pbyr<1965 then yrbirth=13;
if 1965<=pbyr<1970 then yrbirth=14;
if 1970<=pbyr<1975 then yrbirth=15;
if 1975<=pbyr<1980 then yrbirth=16;
if 1980<=pbyr<1985 then yrbirth=17;
if 1985<=pbyr<1990 then yrbirth=18;
if 1990<=pbyr<1995 then yrbirth=19;

Finally, in response to your questions regarding the variables - again, as a courtesy and without any obligation to do so - we are providing copies of the B-Reader Form, a version of the paper questionnaire (the questionnaire was administered by computer in the field), and the data layouts provided by NORC.

Again, I trust this more than answers your questions.

All the best,

Addy R. Schmitt
Assistant United States Attorney
Civil Division
U.S. Attorney's Office for the District of Columbia
501 3rd Street, NW | 4th Floor | Washington, D.C. 20530
202-252-2530 | 202-252-2599 | addy.schmitt@usdoj.gov

Please note the new phone and fax numbers.

APPENDIX C – 9

Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law 106-554)

Sec. 515. (a) In General.--The Director of the Office of Management and Budget shall, by not later than September 30,2001, and with public and Federal agency involvement, issue guidelines under sections 3504(d)(1) and 3516 of title 44, United States Code, that provide policy and procedural guidance to Federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by Federal agencies in fulfillment of the purposes and provisions of chapter 35 of title 44, United States Code, commonly referred to as the Paperwork Reduction Act.

(b) Content of Guidelines.--The guidelines under subsection (a) shall--

(1) apply to the sharing by Federal agencies of, and access to, information disseminated by Federal agencies; and

(2) require that each Federal agency to which the guidelines apply--

(A) issue guidelines ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by the agency, by not later than 1 year after the date of issuance of the guidelines under subsection (a);

(B) establish administrative mechanisms allowing affected persons to seek and obtain correction of information maintained and disseminated by the agency that does not comply with the guidelines issued under subsection (a); and

(C) report periodically to the Director--

(i) the number and nature of complaints received by the agency regarding the accuracy of information disseminated by the agency; and

(ii) how such complaints were handled by the agency.

The full text of Public Law 106-554 is available through the Government Printing Office website.

APPENDIX C – 10

EPA United States Environmental Protection Agency

Newsroom
News Releases By Date

EPA Administrator Gina McCarthy Testimony Before House Committee on Science, Space and Technology

Release Date: 11/14/2013
Contact Information: press@epa.gov

WASHINGTON – As prepared for delivery.

Good morning Chairman Smith, Ranking Member Johnson, and other distinguished members of the Committee. I am pleased to be here to talk about the central role science plays at the U.S. Environmental Protection Agency.

Let me begin by stating that science is and has always been the backbone of the EPA's decision-making. The Agency's ability to pursue its mission to protect human health and the environment depends upon the integrity of the science upon which it relies. I firmly believe that environmental policies, decisions, guidance, and regulations that impact the lives of all Americans must be grounded, at a most fundamental level, in sound, high quality, transparent, science.

Because we rely so heavily on science to meet our mission on behalf of the American people, it must be conducted in ways that are transparent, free from bias and conflicts of interest, and of the highest quality, integrity, and credibility. These qualities are important not just within our own organization and the federal government, but across the scientific community, with its long established and highly honorable commitment to maintaining strict adherence to ethical investigation and research. That's why the agency has established—and embraced—a Scientific Integrity Policy that builds upon existing Agency and government-wide policies and guidance documents, explicitly outlining the EPA's commitment to the highest standards of scientific integrity. And that commitment extends to any scientist or organization who wishes to contribute to our efforts. All EPA-funded research projects, whether conducted by EPA scientists or outside grantees and collaborators, must comply with the agency's rigorous quality assurance requirements.

To ensure that we have the best possible science, we are committed to rigorous, independent peer review of the scientific data, models and analyses that support our decisions. Peer review can take a number of forms, ranging from external reviews by the National Academy of Sciences or the EPA's federal advisory committees to contractor-coordinated reviews. Consistent with OMB guidance, we require peer review for all EPA research products and for all influential scientific information and highly influential scientific assessments.

Among the external advisory committees is the EPA Science Advisory Board (SAB). SAB reviews are conducted by groups of independent non-EPA scientists with the range of expertise required for the particular advisory topic. We invite the public to nominate experts for SAB panels and to comment on candidates being considered by the EPA for SAB panels. The EPA evaluates public comments and information submitted about SAB nominees. The EPA reviews experts' confidential financial information to ensure that there are no conflicts of interest.

SAB peer reviews are conducted in public sessions in compliance with the open-government requirements of the Federal Advisory Committee Act. The public is invited to attend and to provide oral and written comments for consideration by the SAB. Public comments help to ensure that all relevant scientific and technical issues are available to the SAB as it reviews the science that will support our environmental decisions.

Another example is the Clean Air Scientific Advisory Committee (CASAC) which provides independent advice to the EPA Administrator on the science that supports the EPA's National Ambient Air Quality Standards. The CASAC reviews the EPA's Integrated Science Assessments which deliver science in support of the Clean Air Act.

Thanks to the science behind the implementation of the Clean Air Act, we have made significant and far-reaching improvements in the health and well-being of the American public. In 2010 alone, EPA estimates that programs implemented pursuant to the Clean Air Act Amendments of 1990 avoided 160,000 premature deaths millions of cases of respiratory problems such as acute bronchitis and asthma attacks; 45,000 cardiovascular hospitalizations; and 41,000 hospital admissions. These improvements have all occurred during a period of economic growth; between1970 and 2012 the Gross Domestic Product increased by 219 percent.

Through a transparent and open process, we have also committed to enhancing the Agency's Integrated Risk Information System (IRIS) assessment program. A strong, scientifically rigorous IRIS Program is of critical importance, and the EPA is in the process of: 1) enhancing the scientific integrity of assessments; 2) enhancing the productivity of the Program; and 3) increasing transparency so that issues are identified and debated early in the process. In 2009, the EPA made significant enhancements to IRIS by announcing a new 7-step assessment development process. Since that time, the National Research Council (NRC) has made recommendations related to enhancing the development of IRIS assessments. The EPA is making changes to the IRIS Program to implement the NRC recommendations. These changes will help the EPA produce more high quality IRIS assessments each year in a timely and transparent manner to meet the needs of the Agency and the public. A newly released NRC report is largely supportive of the enhanced approach the EPA is taking to develop the IRIS assessment for inorganic arsenic.

Search this collection of releases | or search all news releases

Get news releases by email

View selected historical press releases from 1970 to 1998 in the EPA History website.

Recent additions

12/12/2013 EPA Provides Updated Guidance to Schools on PCB-containing Lighting Fixtures
12/12/2013 Fuel Economy of New Vehicles Sets Record High / Fuel Economy Gains to Continue Under President Obama's Clean Car Programs
12/12/2013 EPA Takes Action to Protect Farm Workers in Puerto Rico; Bayer CropScience to Initiate Measures to Protect Workers and Pay $53,000 Penalty
12/11/2013 EPA Announces Dec. 17 Public Availability Session in York, Neb., to Discuss Groundwater Contamination Issues
12/11/2013 EPA Proposes Pair of Groundwater Contamination Sites in York, Neb., for Addition to Superfund's National Priorities List

As I mentioned in my opening statement, science is the backbone of our decision-making and our work is based on the principles of scientific integrity and transparency that are both expected and deserved by the American people. I am proud of the EPA's research efforts and the sound use of science and technology to fulfill the EPA's mission to protect human health and safeguard the natural environment.

Thank you for the opportunity to testify before you today. I am happy to answer any questions you may have at this time.

R163

Receive our News Releases Automatically by Email

Last updated on Thursday, December 12, 2013 http://yosemite.epa.gov/opa/admpress.nsf/d0cf6618525a9efb85257359003fb69d/201f4594a4b43bad85257c22007ac270!OpenDocument

APPENDIX C – 11

United States Environmental Protection Agency

Office of Research and Development Washington DC 20460

EPA/600/8-90/066F October 1994

EPA

Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry

GLOSSARY

Activity Median Diameter (AMD)
Refers to the median of the distribution of radioactivity, toxicological, or biological activity with respect to particle size.

Acute Exposure
A one-time or short-term exposure with a duration of less than or equal to 24 h.

Aerodynamic Diameter
Term used to describe particles with common inertial properties to avoid the complications associated with the effects of particle size, shape, and physical density.

Aerodynamic Equivalent Diameter (d_{ae})
"Aerodynamic diameter" generally used. The diameter of a unit density sphere ($\rho_p = 1$ g/cm^3) having the same settling velocity (due to gravity) as the particle of interest of whatever shape and density. Refer to Raabe (1976) and Appendix H for discussion.

Aerodynamic (Viscous) Resistance Diameter (d_{ar})
The "Lovelace" definition for aerodynamic diameter. Characteristic expression based on terms describing a particle in the Stokes' regime. Refer to Raabe (1976) for equation.

Aerosol
All-inclusive term. A suspension of liquid or solid particles in air.

ATPS
Ambient temperature and pressure, saturated (a condition under which a gas volume is measured).

BTPS
Body temperature and pressure, saturated (a condition under which a gas volume is measured).

Critical Effect
The first adverse effect, or its known precursor, that occurs as the dose rate increases. Designation is based on evaluation of overall data base.

Chronic Exposure
Multiple exposures occurring over an extended period of time, or a significant fraction of the animal's or the individual's lifetime.

Dosimetric Adjustment Factor (DAF)
A multiplicative factor used to adjust observed experimental or epidemiological data to human equivalent concentration for assumed ambient scenario. See regional gas dose ratio (RGDR) and regional deposited dose ratio (RDDR).

1. INTRODUCTION AND OVERVIEW

This document describes the U.S. Environmental Protection Agency (EPA) methodology for estimation of inhalation reference concentrations (RfCs) (earlier terminology was "inhalation reference dose" or "RfD_i") as benchmark estimates of the quantitative dose-response assessment of chronic noncancer toxicity for individual inhaled chemicals. Noncancer toxicity refers to adverse health effects other than cancer and gene mutations. This overview chapter discusses general principles of dose-response assessment for noncancer toxicity, the development of the RfC methodology, and its role within the context of the risk assessment process. Subsequent chapters of the document discuss criteria and information to be considered in selecting key studies for RfC derivation, provide an overview of the respiratory system and its intra- and interspecies variables, and discuss areas of uncertainty and data gaps in relation to the proposed methodology.

1.1 INHALATION REFERENCE CONCENTRATION: DEVELOPMENT, DEFINITION, AND DERIVATION

The EPA has a history of advocating the evaluation of scientific data and calculation of Acceptable Daily Intake (ADI) values for noncarcinogens as benchmark values for deriving regulatory levels to protect exposed populations from adverse effects. For example, the Office of Pesticide Programs has long used the concept of ADI for tolerance estimates of pesticides in foodstuffs, the Office of Health and Environmental Assessment (OHEA) has used ADI values for characterizing levels of pollutants in ambient waters (Federal Register, 1980), and the National Research Council (1977, 1980) has recommended the ADI approach to characterize levels of pollutants in drinking water with respect to human health.

In 1983, the National Academy of Sciences (NAS) published a report entitled "Risk Assessment in the Federal Government: Managing the Process" (National Research Council, 1983). The NAS had been charged with evaluating the process of risk assessment as performed at the federal level in order to determine the "mechanisms to ensure that government regulation rests on the best available scientific knowledge and to preserve the integrity of scientific data and

judgements" so that controversial decisions regulating chronic health hazards could be avoided. The NAS recommended that the scientific aspects of risk assessment should be explicitly separated from the policy aspects of risk management. Risk assessment, as shown in Figure 1-1, was defined as the characterization of the potential adverse human health effects of exposures to environmental hazards and consists of the following four steps: (1) hazard identification: the determination of whether a chemical is or is not causally linked to a particular health effect; (2) dose-response assessment: the estimation of the relation between the magnitude of exposure and the occurrence of the health effects in question; (3) exposure assessment: the determination of the extent of human exposure; and (4) risk characterization: the description of the nature and often the magnitude of human risk, including attendant uncertainty.

Following the NAS report, the EPA developed a methodology for evaluating available data pertaining to xenobiotics for purposes of developing oral reference doses (RfDs) (Barnes and Dourson, 1988). Although similar to ADIs in intent, RfDs were based upon a more rigorously defined methodology that adhered to the principles proposed by the NAS and included guidance on the consistent application of uncertainty factors for prescribed areas of extrapolation required in the operational derivation. The RfD methodology represents a quantitative approach to assess toxicity data in order to derive a dose-response estimate. According to the NAS paradigm, the final step of the risk assessment process, risk characterization, would involve the comparison of the RfD as a dose-response estimate with an exposure estimate.

The RfC methodology to estimate benchmark values for noncancer toxicity of inhaled chemicals significantly departed from the RfD approach. The same general principles were used, but the RfC methodology was expanded to account for the dynamics of the respiratory system as the portal of entry. The major difference between the two approaches, therefore, is that the RfC methodology includes dosimetric adjustments to account for the species-specific relationships of exposure concentrations to deposited/delivered doses. The physicochemical characteristics of the inhaled agent are considered as key determinants to its interaction with the respiratory tract and ultimate disposition. Particles and gases are treated separately, and the type of toxicity observed (respiratory tract or toxicity remote to the portal-of-entry) influences the dosimetric adjustment applied.

An inhalation reference concentration (RfC) is defined as an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human

exposure-dose-response continuum and will therefore be revised accordingly, it must be recognized that the definition of HEC is iterative and dynamic as well. That is, the HEC is a concentration back-extrapolated from an appropriate surrogate internal dose to the extent that this has been defined.

Although it is preferable to use human studies as the basis for the dose-response derivation, adequate human data are not always available, often forcing reliance on laboratory animal data. Presented with data from several animal studies, the risk assessor first seeks to identify the animal model that is most relevant to humans, based on comparability of biological effects using the most defensible biological rationale; for instance, by using comparative metabolic, pharmacokinetic, and pharmacodynamic data. In the absence of a clearly most relevant species, however, the most sensitive species is used as a matter of science policy at the EPA. For RfCs, the most sensitive species is designated as the species that shows the critical adverse effect at an exposure level that, when dosimetrically adjusted, results in the lowest HEC.

The critical toxic effect used in the dose-response assessment is generally characterized by the lowest $NOAEL_{[HEC]}$ that is also representative of the threshold region (the region where toxicity is apparent from the available data) for the data array. The objective is to select a prominent toxic effect that is pertinent to the chemical's key mechanism of action. This approach is based, in part, on the assumption that if the critical toxic effect is prevented, then all toxic effects are prevented (see Section 1.2, general principles of dose-response assessment for noncancer toxicity). The determination of the critical toxic effect from all effects in the data array requires toxicologic judgment because a chemical may elicit more than one toxic effect (endpoint) in tests of the same or different exposure duration, even in one test species. Further, as discussed in Appendix A, the NOAEL and LOAEL obtained from studies depend on the number of animals or subjects examined and on the spacing of the exposure levels. The $NOAEL_{[HEC]}$ from an individual study (or studies) that is also representative of the threshold region for the overall data array is the key datum synthesized from an evaluation of the dose-response data. Determination of this critical effect represents the first scientific evaluation required by the RfC dose-response assessment.

The RfC is an estimate that is derived from the $NOAEL_{[HEC]}$ for the critical effect by consistent application of uncertainty factors (UFs). The UFs are applied to account for

2. QUALITATIVE EVALUATION OF THE DATA BASE

This chapter outlines considerations for the collection and qualitative evaluation of diverse data into a cohesive toxicity profile that then can be evaluated by means of the quantitative procedures for dose-response analysis provided in Chapter 4. The conceptual basis for the dosimetry adjustments applied to inhaled agents and other considerations specific to this administration route are addressed in Chapter 3.

The aim of the inhalation reference concentration (RfC) methodology is to establish a relationship between a particular agent in the air and a specific health effect (or effects). To define such a relationship, evidence must be collected from diverse sources and synthesized into an overall judgment of health hazard (Hackney and Linn, 1979). One of the major challenges to performing dose-response assessment for noncancer endpoints is that it requires the evaluation of effects measured in a number of different tissues. Often different endpoints are investigated in different studies, in different species, and at various concentrations. The effects measured may represent different degrees of severity (adversity) within disease continuums. Qualitative evaluation of the data base, also known as the hazard identification component of risk assessment, involves integrating a diverse array of data into a cohesive, biologically plausible toxicity "picture" or weight-of-the-evidence relationship to establish that the agent causes an effect (or effects) and is of potential human hazard. Questions addressed by this process include whether the agent associated with an effect is responsible for the effect, if the effect is biologically significant, and what the potential public health implications might be. Answering such questions requires ascertaining the validity and meaning of the toxicity data, determining whether the experimental results as a whole suggest or show causality between the agent and the effect, and evaluating whether or not the causal relationship is applicable under other sets of circumstances (e.g., in extrapolating from test animals to humans). This entails consideration of all relevant human and laboratory animal data of various study types, studies with differing results (e.g., positive and negative), pharmacokinetic disposition data (deposition, absorption, distribution, metabolism, elimination) mechanistic information, and structure-activity relationships. This process integrates information needed for the dose-response assessment, which is discussed in

TABLE 2-3. COMPARISON OF THE QUALITIES OF FIELD AND EXPERIMENTAL APPROACHES IN THE STUDY OF THRESHOLD LIMIT VALUE/BIOLOGIC EXPOSURE INDICES RELATIONSHIPS

	Approach	
Factor	Field	Experimental
Exposure (dose) measurement	+ +	+ + +
Physical workload characterization	+	+ + +
Timing of biological sampling	+	+ + +
Effects of exposure repetition	+ + +	+ +
Environmental variability	+ +	+ +
Representativity of the subjects	+ + +	+

+++ = Good; ++ = Medium; + = Poor.

Source: Droz (1985).

Application of Physiologically Based Pharmacokinetic Models

Physiologically based pharmacokinetic models are simulation models described by simultaneous differential equations, the number of which is dictated by the number of compartments needed to describe the physiological and metabolic processes involved. In the context of characterizing the exposure-dose-disease continuum, simulation models can be considered as complementary, providing critical insight on key processes related to the fate of chemicals in the body and for depicting the contribution of various exposure and biological factors to the variability of response. That is, these models can provide the following information on which biological monitoring (e.g., BEIs) is designed and data are interpreted: (1) concentration-effect relationships, (2) time-effect relationships, (3) matching exposure in the workplace with integrated exposure, (4) depicting effects of external and internal factors that alter the relationship between intensity of exposure and biological concentration and body burden of the biologic marker, (5) extrapolation and prediction of biological concentrations resulting from exposure to new compounds or new exposure conditions, and (6) verification of data (Leung, 1992; Fiserova-Bergerova, 1990; Leung and Paustenbach, 1988; Droz, 1985). Simulation models, because of their ability to match the extent of exposures associated with the predetermined dose or biological markers of exposure, are a valuable tool

in extrapolation of reference values for workers with unusual workshifts (Andersen et al., 1987b; Saltzman, 1988).

2.1.1.2 Epidemiologic Data

There are essentially three areas of concern in assessing the quality of an epidemiologic study. These involve the design and methodological approaches used for: (1) exposure measures, (2) effect measures, and (3) the control of covariables and confounding variables (Lebowitz, 1983). The study population and study design must adequately address the health effect in question in order to support a risk assessment (Lebowitz, 1983). In order to accomplish this goal, the exposure measures must be appropriate and of sufficient quality; the statistical analysis methods must be suitable to the study design and goals; the health effect measures must be reliable and valid; and the covariables and confounding variables need to be controlled or eliminated. Additional guidance on evaluation of the quality of individual epidemiologic studies is provided in Appendix B. Criteria for causal significance are provided in Appendix C.

Assessment of Exposure Measures

The problem of the accuracy and relevance of exposure measurements is not unique to epidemiologic investigations, but it can be exacerbated due to the long-term nature of these studies. For example, the nature of aerometric data may change over time because of different air sampling techniques. Exposures also change over time because of different industrial hygiene practices and because individuals change jobs and residences. Accurate documentation of air toxicant levels, therefore, is critical in determining the usefulness of an investigation as well as documentation that the analysis of the air toxicant is appropriate and of sufficient sensitivity. It also is advisable to have the concentrations of other pollutants reported and considered in the statistical analyses to help rule out confounding or interactive effects. The number, location, and timing of monitors should be suitable to allow an appropriate determination of exposure of the subjects to the pollutant being studied and to the pollutants that could confound the results. When appropriate, the exposure measure or estimate should take into account indoor/outdoor exposures and activity and subject location data. Unfortunately, exposure measures often are the weakest component of an

Other considerations include the adequacy of study duration and quality of the follow-up. A disease with a long latency before clinical presentation requires a longer study duration than one with an acute onset. Valid ascertainment (such as verification according to the International Classification of Diseases IX) of the causes of morbidity and death also is necessary.

Evaluation of epidemiologic studies may require interpretation of a variety of subjective health effects data. Questionnaire responses may be biased by the way questions are worded, the training of an interviewer, or the setting. However, a study based on a high-quality questionnaire can provide useful results. For example, a committee of the American Thoracic Society (ATS) charged with defining an adverse respiratory health effect, has come to a consensus that "in general, increased prevalence of chronic respiratory symptoms as determined from questionnaire surveys should be considered to be an adverse health effect" (American Thoracic Society, 1985). Questionnaires should be validated as part of the investigation protocol, unless a standard questionnaire that has previously been validated is used (Medical Research Council, 1960; Ferris, 1978; National Institute for Occupational Safety and Health, 1986).

It is very important to consider differences between statistical significance and medical or biological significance. Both the variability of an outcome measure and the magnitude of an exposure's effect determine the level of statistical significance. For example, data from a large study population analyzed with sophisticated techniques may yield statistically significant effects of small magnitude that cannot readily be interpreted biologically. Conversely, apparently large changes of clinical importance may not be statistically significant if the study population is too small. In addition, some studies present false negative or no-effect results due to the lack of power. Judgments concerning medical or biological significance should be based on the magnitude and class of a particular effect. For example, cough or phlegm production can be considered less important than effects resulting in hospital admissions, but daily productive cough can be more important than infrequent cough. Underlying assumptions and nuances of the statistical procedures applied to the data also need to be considered. This will probably best be accomplished on a case-by-case basis.

Because the RfC considers both portal-of-entry and remote (systemic) effects, it would be helpful to define an "adverse respiratory health effect." An ATS committee published

guidelines that defined such an effect as medically significant physiologic or pathologic changes generally evidenced by one or more of the following (American Thoracic Society, 1985):

- Interference with the normal activity of the affected person or persons
- Episodic respiratory illness
- Incapacitating illness
- Permanent respiratory injury or
- Progressive respiratory dysfunction

Appendix D provides detailed descriptions of adverse respiratory effects in humans.

Assessing the Control of Confounding and Covariables

Epidemiologic investigations attempt to relate an exposure to a given health effect, but this includes accounting for the "background" health effect (pathologic condition) that exists in individuals due to predisposing factors and preexisting health conditions, or from other variables, such as occupational exposures.

Various host factors contribute as risk factors for disease and can influence the health indices assessed. For example, asthmatics may be particularly susceptible to effects from exposure to irritant gases. Epidemiologic evaluation of these factors often not only accounts for such interactions but also can help to characterize susceptible or sensitive groups. Covariables can be as important as the major aerometric variables themselves in affecting human health. Other exposures, such as concomitant occupational exposures and smoking, in particular, can affect the disease outcome. Meteorologic variables such as air velocity, temperature, and humidity also are very important factors when considering respiratory health effects. These covariables should be controlled by both the study design and analysis, as appropriate.

The final step in the inferential process from an epidemiologic investigation is the extension of the study results to persons, populations, or settings not specifically included in the experimental design, that is, to demonstrate consistency of results within replicates in

2.1.2.4 Study Validity and Relevance to Extrapolation

The validity of the study and its relevance to human extrapolation is another major area to consider when assessing individual animal studies. It involves the evaluation of a number of factors, including all elements of exposure definition (concentration, duration, frequency, administration route, and physicochemical characterization of the chemical used), reliability of and limits to the procedures used for both exposure and effects measurements, relevance of the exposure level tested to the anticipated human exposure level, nature of the effect (consistency with the area of toxicology assessed and the suspected mechanism of action), and the similarities and differences between the test species and humans (e.g., in absorption and metabolism).

Animal studies are conducted using a variety of exposure scenarios in which the concentration, frequency, and duration of exposure may vary considerably. Studies may use different durations (acute, subchronic, and chronic) as well as schedules (single, intermittent, and continuous). All of these studies contribute to the hazard identification of the risk assessment. Special consideration should be addressed to those studies of appropriate duration for the reference level to be determined (i.e., chronic investigations for the RfC).

These exposure concerns (concentration and duration) are compounded when the risk assessor is presented with data from several animal studies. An attempt to identify the animal model most relevant to humans should be made on the most defensible biological rationale (e.g., comparable metabolism and pharmacokinetic profiles). In the absence of such a model, the most sensitive species (i.e., the species showing a toxic effect at the lowest administered dose) is adopted for use as a matter of science policy at the EPA (Barnes and Dourson, 1988). This selection process is more difficult if the laboratory animal data are for various exposure routes, especially if the routes are different from that in the human situation of concern.

Because the data base may be deficient for the route of exposure of interest, it is the EPA's view that the toxicity potential manifested by one route can be indicative of potential toxicity via any other exposure route unless convincing contrary evidence exists (Barnes and Dourson, 1988). Quantitative extrapolation, however, requires consideration of the differences in the dosimetry for the chemical resulting from the different exposure routes. Detailed consideration is given to route-to-route extrapolation in Section 4.1.2.

design, or is a function of designating a specified health effect measure (e.g., 10% incidence of a lesion) as the outcome of interest in the case of some alternative approaches presented in Appendix A[3], and therefore, does not necessarily reflect the "true" biological threshold.

Table 4-2 presents the four types of effect levels that may be applicable when evaluating an individual study. Historically, the distinction between adverse effects and nonadverse effects has been and remains problematic. For example, although disease is a dynamic process (injury, adaptation, or healing), a pathologist records a morphologic change at a single point in time and these "freeze-frame" data are used to determine the probable cause and pathogenesis (past) and probable progression or outcome (future). Designation of an effect level (i.e., the designation of adversity) requires interpretation of the data based on an ability to deduce the preceding events that have led to the observed change and to predict the outcome or progression. The relationship between structural alterations to altered function is not always simple, however.

Determining whether altered morphology is an adaptive response or truly an expression of toxicity (functional impairment) can be extremely difficult and even controversial (Burger et al., 1989; Ruben and Rousseaux, 1991). In some cases, structural alteration can occur, but normal function can continue in target tissues with functional reserve such as the lung, liver, and kidney. Not all tissues demonstrate this high reserve. The central nervous system can compensate to only a limited degree and where the damage occurs is vitally important for the function of the system. Therefore, "focal" damage may be adverse in some but not all target tissues. Also, the lack of observed functional change may be due to failure to detect subtle or unknown functional changes rather than to their absence.

A similar morphologic alteration may have both functional and physiologic significance, but often it is difficult to differentiate toxicity from physiologic response by morphologic means alone. Not all functional abnormalities manifest themselves morphologically. Temporal-spatial patterns are particularly challenging when evaluating toxicologic pathology. Problems concerning time include reversibility, adaptation versus toxicity, progression versus

---

[3]There are alternative approaches under development (presented and discussed in Appendix A) aimed at deriving estimates of exposures that are analogous in intent to the establishment of a NOAEL. The NOAEL/LOAEL approach outlined is not intended to discourage alternative or more sophisticated dose-response procedures when sufficient data are available, but rather to present key issues necessarily involved (e.g., dosimetric adjustment and data array analysis) in any approach for the assessment of noncancer toxicity.

TABLE 4-2. FOUR TYPES OF EFFECT LEVELS[a] (RANKED IN ORDER OF INCREASING SEVERITY OF TOXIC EFFECT) CONSIDERED IN DERIVING INHALATION REFERENCE CONCENTRATIONS FOR NONCANCER TOXICITY

NOEL:	No-Observed-Effect Level. That exposure level at which there are no statistically and biologically significant increases in frequency or severity of effects between the exposed population and its appropriate control.
NOAEL:	No-Observed-Adverse-Effect Level. That exposure level at which there are no statistically and biologically significant increases in frequency or severity of adverse effects[b] between the exposed population and its appropriate control. Effects are produced at this level, but they are not considered to be adverse.
LOAEL:	Lowest-Observed-Adverse-Effect Level. The lowest exposure level in a study or group of studies that produces statistically and biologically significant increases in frequency or severity of adverse effects between the exposed population and its appropriate control.
FEL:	Frank Effect Level[c]. That exposure level that produces frankly apparent and unmistakable adverse effects, such as irreversible functional impairment or mortality, at a statistically and biologically significant increase in frequency or severity between an exposed population and its appropriate control.

[a]Note that these levels represent points on a continuum and are not discrete.
[b]Adverse effects are defined as any effects resulting in functional impairment and/or pathological lesions that may affect the performance of the whole organism, or that reduce an organism's ability to cope with an additional challenge.
[c]Frank effects are defined as overt or gross adverse effects (e.g., severe convulsions, lethality, etc.).

regression, and peracute lethal toxicity. Problems concerning space are limited to missing the lesion completely or missing a relevant area because of sampling method. For example, histologic examination of the nasal cavity should select four tissue sections, not one, to achieve a thorough examination (Young, 1981). Further, due to the proximal to distal inspiratory airstream, some examination of the upper respiratory tract is indicated when respiratory toxicity from an inhaled irritant is evident in the lower respiratory tract.

Due to the structural-functional and temporal-spatial problems discussed above, an approach that integrates pathological studies (ultrastructural, histochemical, cellular, and molecular) with functional methods is recommended (Ruben and Rousseaux, 1991). Morgan (1991) has provided guidance on the identification and interpretation of URT lesions in toxicologic studies. A systematic but flexible approach to evaluation of lesions in the URT is

recommended, one that considers selection of section level in context with the physicochemical characteristics of the inhaled gas (e.g., water solubility and reactivity), the role of factors that may account for lesion distribution (e.g., dosimetry and tissue susceptibility), and development of a pathogenesis profile or a chronological order of events (e.g., degenerative, adaptive, and adaptive/regenerative changes versus time). The nasal diagrams proposed by Mery et al. (in press) offer an approach to recording data and mapping lesions that aids this type of interpretation strategy. This approach is also likely the best to compile the data and precludes the restraint to interpretation and mathematical modeling presented by data scored categorically for severity (e.g., + = mild, ++ = moderate; and +++ = severe) and/or without sufficient section detail with respect to lesion location (Jarabek, 1994).

In the early stages of respiratory disease, there is considerable uncertainty concerning how to differentiate between acute reversible effects, which are the immediate consequence of an exposure episode, and potential progression to chronic, nonreversible respiratory pathology. The boundary between adaptive and toxic responses also remains controversial for some respiratory tract lesions (Burger et al., 1989). These are important issues both in terms of evaluation of respiratory tract effects per se, as well as for decisions concerning the critical effect in inhalation studies. Inhalation-specific issues such as evaluation of pulmonary function, sensory irritation, and allergic sensitization data are discussed in Section 2.2.

Designation of effect levels usually contains an element of scientific judgment in addition to objective criteria. Considerable experience and precedent for such decisions have accrued over the last several years in the process of developing oral reference doses, RfCs, and other health-related benchmark estimates. Table 4-3 presents guidance as to how general effects would usually be designated as different (adverse) effect levels. In general, effects that may be considered marginal are designated as adverse only to the extent that they are consistent with other structural and functional data suggesting the same toxicity. For example, altered liver enzymes (statistically out of normal range) would only be considered adverse in context with altered structure (pathology) and liver weight changes.

APPENDIX C – 12

EXCERPTS

EPA-540-R-070-002
OSWER 9285.7-82
January 2009

Risk Assessment Guidance for Superfund
Volume I: Human Health Evaluation Manual
(Part F, Supplemental Guidance for Inhalation Risk Assessment)

Final

Office of Superfund Remediation and Technology Innovation
Environmental Protection Agency
Washington, D.C.

Category 3 gases are relatively water-insoluble and are unreactive in the respiratory tract (e.g., benzene, styrene). Their toxicity is generally at sites remote to the respiratory tract (USEPA, 1994). The DAF for Category 3 gases is based on the ratio of the animal blood:gas partition coefficient ($H_{b/g\text{-}animal}$) and the human blood:gas partition coefficient ($H_{b/g\text{-}human}$). See Appendix A, Section 4 of this guidance for an example of a Category 3 DAF equation.

Category 2 gases are moderately water-soluble and may be rapidly reversibly reactive or moderately to slowly irreversibly reactive in respiratory tract tissue (e.g., acetonitrile, xylene, propanol, isoamyl alcohol). These gases have potential for significant accumulation in the blood, so they can exhibit both respiratory and remote toxicity (USEPA, 1994). The DAF for respiratory effects of Category 2 gases consists of an RGDR and is based on the animal to human ratio of the V_e and the SA of the region of the respiratory tract where the effect occurs, as for Category 1 gases. The DAF for extra-respiratory (ER) effects of a Category 2 gas is based on the ratio of the $H_{b/g\text{-}animal}$ and the $H_{b/g\text{-}human}$, as for Category 3 gases.

Particles also vary by solubility and reactivity. However, the default equations used to estimate the predicted regional deposition fractions for particles are based on non-soluble, non-hygroscopic particles (USEPA, 1994, Section 4.3.5.3). The DAF for a particle causing an effect in the respiratory tract is the $RDDR_r$. The $RDDR_r$ is based on the animal to human ratio of the V_e and the fractional deposition of the particle in that region (F_r), divided by the SA_r of the region where the effect occurs. This derivation, from the *Inhalation Dosimetry Methodology*, conservatively assumes that 100 percent of the deposited dose remains in the respiratory tract; clearance mechanisms are not considered. The DAF for a particle causing an ER effect, the $RDDR_{ER}$, is based on the animal to human ratio of the V_e and the total deposition of the particle in the entire respiratory tract (F_{total}), divided by BW (USEPA, 1994). The $RDDR_{ER}$ assumes that 100 percent of the deposited dose in the entire respiratory tract is available for uptake into the systemic circulation. See Appendix A, Section 5 for examples of specific particle DAF equations.

2.1.2 Default Approach - Extrapolation from Human Occupational Data

When human data are available to derive an RfC, duration adjustments are often required to account for differences in exposure scenarios (e.g., extrapolation from an 8 hour/day occupational exposure to a continuous chronic exposure). The default approach recommended by the *Inhalation Dosimetry Methodology* for adjusting the POD concentration (e.g., the no observable adverse effect level (NOAEL)) obtained from human study data is provided below in Equation 3 (USEPA, 1994, Equation 4-49).[17,18]

[17] If sufficient data are available, a PBPK model or intermediate approach using chemical-specific information may be employed in preference to the default method for extrapolating human occupational data to an HEC.

[18] EPA's IRIS glossary defines an adverse effect as the following: "A biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to respond to an additional environmental challenge" (USEPA, 2008b).

$NOAEL_{[HEC]}$ = NOAEL x (VEho/VEh) x 5 days/7 days (Equation 3)

Where: $NOAEL_{[HEC]}$ (mg/m^3) = the NOAEL or analogous exposure level obtained with an alternate approach, dosimetrically adjusted to an ambient HEC;
NOAEL (mg/m^3) = occupational exposure level (time-weighted average over an 8-hour exposure period);
VEho = human occupational default minute volume over 8 hours (10 m^3); and
VEh = human ambient default minute volume over 24 hours (20 m^3).

2.2 Derivation of the Inhalation Unit Risk

The default approach for determining predictive cancer risk recommended by EPA's *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005a; hereafter, *Cancer Guidelines*) is a linear extrapolation from exposures observed in the animal or human occupational study.[19] This approach involves drawing a straight line from the POD to the origin. The default linear extrapolation approach is generally considered to be conservatively protective of public health, including sensitive sub-populations (USEPA, 2005a). The slope of this line is commonly called the slope factor, and when the units are risk per $\mu g/m^3$, it is also called the IUR. EPA defines an IUR in the IRIS glossary as "the upper-bound excess lifetime cancer risk estimated to result from continuous exposure to an agent at a concentration of 1 $\mu g/m^3$ in air" (USEPA, 2008b). Equation 4 below presents a linear extrapolation from a POD of 10 percent response (LEC_{10}).[20]

$IUR = 0.1/LEC_{10[HEC]}$ (Equation 4)

Where: $IUR\ (\mu g/m^3)^{-1}$ = Inhalation Unit Risk; and
$LEC_{10[HEC]}$ ($\mu g/m^3$) = the lowest effective concentration using a 10 percent response level, dosimetrically adjusted to an HEC.

2.3 Derivation of the Reference Concentration

EPA defines an RfC in the IRIS glossary as "an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious effects during a lifetime" (USEPA, 2008b). The RfC is derived after a review of the health effects database for a chemical and identification of the most sensitive and relevant endpoint along with the principal study or studies demonstrating that endpoint. EPA Chemical Managers use UFs to account for recognized

[19] According to the *Cancer Guidelines*, "[a] nonlinear approach should be selected when there are sufficient data to ascertain the mode of action [MOA] and conclude that it is not linear at low doses and the agent does not demonstrate mutagenic or other activity consistent with linearity at low doses" (USEPA, 2005a, page 3-22). In addition, [l]inear extrapolation should be used when there are MOA data to indicate that the dose-response curve is expected to have a linear component below the POD" (USEPA, 2005a, page 3-21). This information will appear on the IRIS profile or other toxicological information source for a chemical. Chemicals with a mutagenic MOA are thought to pose a higher risk during early life. Procedures for assessing cancer risk from these chemicals are outlined in Section 5.1.

[20] The POD used in Equation 4 is an LEC_{10}, which is the lower 95 percent confidence limit on the concentration corresponding to a 10 percent response rate (i.e., the EC_{10}). Other PODs may be substituted for this value, which could be associated with alternative response levels (e.g., 1 percent, 5 percent).

uncertainties in the extrapolations from the experimental data conditions to an estimate appropriate to the assumed human scenario (USEPA, 1994). See Table 3 for a description of the standard UFs. The formula used for deriving the RfC from the HEC is provided below.

$$RfC = NOAEL_{[HEC]}/(UF)^{1} \qquad \text{(Equation 5)}$$

Where:

RfC (mg/m^3) = Reference Concentration

$NOAEL_{[HEC]}$ (mg/m^3) = The NOAEL or analogous exposure level obtained with an alternate approach, dosimetrically adjusted to an HEC; and

UF = Uncertainty factor(s) applied to account for the extrapolations required from the characteristics of the experimental regimen.

[1] Some toxicological information sources for RfCs will incorporate an additional factor to account for deficiencies in the available data set, called a modifying factor (MF). In 2002, however, EPA published the *RfD/RfC Review*, which recommended that the use of MFs be discontinued because their purpose is "sufficiently subsumed in the general database UF" (USEPA, 2002c, page xviii). Therefore, RfCs published subsequent to this document will not include MFs.

APPENDIX C – 13

EXCERPTS

EPA/630/P-02/002F
December 2002
Final Report

A REVIEW OF THE REFERENCE DOSE AND REFERENCE CONCENTRATION PROCESSES

Prepared for the
Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC

Reference Dose/Reference Concentration (RfD/RfC) Technical Panel

Bob Benson (OPRA/Region 8)
Gary Foureman (NCEA/ORD)
Lee Hofmann (PARMS/OSWER)
Carole Kimmel (NCEA/ORD)*
Gary Kimmel (NCEA/ORD)
Susan Makris (OPP/OPPTS)
Deirdre Murphy (OAQPS/OAR)
Edward Ohanian (OST/OW)
Jennifer Orme-Zavaleta (NHEERL/ORD)
Deborah Rice (NCEA/ORD)
Jennifer Seed (OPPT/OPPTS)
Hugh Tilson (NHEERL/ORD)
Vanessa Vu (SAB Staff Office, formerly OSCP/OPPTS and NCEA/ORD)

*Technical Panel Chair

Technical Advisors

Amy Mills, IRIS Director, NCEA/ORD
Bill Wood, RAF Director, NCEA/ORD

Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC 20460

Professional judgment is required to decide, on the basis of a thorough review of all available data and studies, whether any observed effect is adverse and how the results fit with what is known about the underlying mode of action. These judgments require the input of experts trained in toxicology, statistics, and epidemiology and, often, of specialists in the structure and function of the target organ systems. Both the biological and the statistical significance of the effects are considered when making these judgments. Biological significance is the determination that the observed effect (a biochemical change, a functional impairment, or a pathological lesion) is likely to impair the performance or reduce the ability of an individual to function or to respond to additional challenge from the agent. Biological significance is also attributed to effects that are consistent with steps in a known mode of action. Statistical significance quantifies the likelihood that the observed effect is not due to chance alone. Precedence is given to biological significance, and a statistically significant change that lacks biological significance is not considered an adverse response.

For many discrete or quantal endpoints (e.g., birth defects, tumors, or some discrete pathological changes), this judgment is more straightforward because criteria have been established for deciding what type and incidence of effects are to be considered to be adverse, and an increase above the background rate can be judged using statistical tools. In the case of continuous measures (e.g., body weight, enzyme changes, physiological measures), this tends to be more difficult, because the amount of change to be considered adverse has not been defined by toxicologists or health scientists. Consequently, the endpoint is often decided in the context of the endpoint itself, the study, and the relationship of changes in that endpoint to other effects of the agent.

Decisions about the amount of change to consider adverse must always be made using professional judgment and must be viewed in light of all the data available on the endpoint of concern. All toxicological data on a chemical must be reviewed before deciding whether an effect is biologically significant and adverse. Using a default cutoff value to define adversity for continuous measures may result in an inappropriate interpretation of data and less than optimum evaluation of a chemical's effects.

4.3.2. Issues to be Considered in Characterizing the Database for Risk Assessment

4.3.2.1. *The Weight-of-Evidence Approach*

A weight-of-evidence approach such as that provided in EPA's RfC Methodology (U.S. EPA, 1994) or in EPA's proposed guidelines for carcinogen risk assessment (U.S. EPA, 1999a) should be used in assessing the database for an agent. This approach requires a critical

evaluation of the entire body of available data for consistency and biological plausibility. Potentially relevant studies should be judged for quality and studies of high quality given much more weight than those of lower quality. When both epidemiological and experimental data are available, similarity of effects between humans and animals is given more weight. If the mechanism or mode of action is well characterized, this information is used in the interpretation of observed effects in either human or animal studies. Weight of evidence is not to be interpreted as simply tallying the number of positive and negative studies, nor does it imply an averaging of the doses or exposures identified in individual studies that may be suitable as PODs for risk assessment. The study or studies used for the POD are identified by an informed and expert evaluation of all the available evidence.

4.3.2.2. *Use of Human and Animal Data in Risk Assessment*

Adequate human data are the most relevant for assessing risks to humans. When sufficient human data are available to describe the exposure-response relationship for an adverse outcome(s) that is judged to be the most sensitive effect(s), reference values should be based on human data. Much more data on a wide range of endpoints typically are required to establish confidence that there are no effects of exposure. If sufficient human data are not available to provide the basis for reference values, data from animal studies must be employed. It is advantageous if some human data are available to compare with effects observed in animals, even if the human data are not adequate for quantitative analysis. Availability of data on effects in humans at least allows qualitative comparison with effects observed in animals for determining whether toxicity occurs in the same organ systems and whether the nature of the effects is similar or different. If no human data are available, reliance must be exclusively on animal data. In that case, attention should be paid to whether data are available in more than one species and, if so, whether the same or similar effects occur in different species and possible sources of any observed differences.

One of the major default assumptions in EPA's risk assessment guidelines is that animal data are relevant for humans (e.g., U.S. EPA, 1991, 1996, 1998c). Such defaults are intended to be used in the absence of experimental data that can provide direct information on the relevance of animal data.

Several types of information should be considered when determining the relevance or nonrelevance of effects observed in animal models for humans. This information is used in a variety of ways, from determining the role of metabolism in toxicity (Is the parent chemical or a metabolite responsible for toxicity?), to assessing whether homologous activity would be

APPENDIX C – 14

EXCERPTS

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

January 30, 2013

EPA-SAB-13-001

The Honorable Lisa P. Jackson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, DC 20460

Subject: Review of EPA's Draft Assessment entitled *Toxicological Review of Libby Amphibole Asbestos* (August 2011)

Dear Administrator Jackson:

EPA's Office of Research and Development (ORD) requested the Science Advisory Board (SAB) to conduct a peer review of EPA's draft Integrated Risk Information System (IRIS) assessment, entitled *Toxicological Review of Libby Amphibole Asbestos (August 2011).* The draft document is the first IRIS assessment specific to Libby Amphibole asbestos (LAA), a term used to refer to the mixture of amphibole mineral fibers identified in the Rainy Creek complex near Libby, Montana. The SAB was asked to comment on the scientific soundness of the hazard and dose-response assessment of LAA-induced cancer and non-cancer health effects.

The SAB finds the EPA's draft assessment to be comprehensive and generally clear, logical and well-written. There are several areas that need more consideration, and we provide recommendations to further enhance the clarity and strengthen the scientific basis for the conclusions presented. The SAB responses to the EPA's charge questions are detailed in the enclosed report. The SAB's major comments and recommendations are provided below:

- Localized pleural thickening is an appropriate health endpoint for the derivation of the inhalation reference concentration (RfC). It is an irreversible structural, pathological alteration of the pleura and is generally associated with reduced lung function. The SAB has identified additional references and recommends that the agency include a more detailed review of the literature to further support this conclusion.

- The SAB supports the derivation of an RfC for LAA based on radiographic evidence of localized pleural thickening in an occupationally exposed Marysville, Ohio, cohort. However, the SAB recommends that the EPA conduct additional analyses to substantiate the RfC (to the extent data permit) of pleural abnormalities using the recently published studies on two other cohorts.

- The SAB recommends that more justification be provided for the selection of the "best" model for non-cancer exposure-response analysis. The SAB also recommends examining other exposure metrics besides the simple cumulative exposure, such as time-weighting of exposures. In addition, more justification is needed for the selection of 10 percent extra risk as the benchmark response since it is not consistent with the guideline for epidemiological data in EPA's *Benchmark Dose Technical Guidance*.

- A composite uncertainty factor of 100 was applied to the point of departure to obtain the RfC. EPA applied an uncertainty factor of 10 to account for human variability and sensitive subpopulations, and a database uncertainty factor of 10 to account for database deficiencies in the available literature for the health effects of LAA. The SAB recommends that the EPA re-evaluate the use of a default database uncertainty factor of 10 as part of the consideration of additional studies; additional data (e.g., Minnesota cohort and data on other amphiboles) might support a lower value, such as 3, for the database uncertainty factor. In addition, the SAB recommends EPA re-visit its judgement of a subchronic-to-chronic uncertainty factor and a LOAEL-to-NOAEL uncertainty factor of 1-fold.

- The SAB agrees that the weight of evidence for LAA supports the descriptor "Carcinogenic to Humans by the Inhalation Route" in accordance with EPA's *Guidelines for Carcinogen Risk Assessment*. The SAB views the mode of carcinogenic action of LAA as complex, and recommends that the agency conduct a formal mode of action analysis in accordance with EPA's *Guidelines for Carcinogen Risk Assessment*. Based on this formal analysis, the agency may still conclude that the default linear extrapolation at low doses is appropriate.

- The SAB supports the selection of the Libby worker cohort for the derivation of the inhalation unit risk (IUR) and agrees that the use of the subcohort post-1959 for quantification may be reasonable due to the lack of exposure information for many of the workers in earlier years. The SAB has suggested sensitivity analyses that would explore the implications of the selection of the subcohort. The SAB finds it appropriate to use lung cancer and mesothelioma as endpoints for the derivation of the IUR. The SAB recommends a more detailed discussion and justification of how the use of mortality data rather than incidence data may have resulted in an undercount of cases of lung cancer and mesothelioma and what implications, if any, it may have for the derivation of the IUR.

- The draft assessment clearly described the methods selected to conduct the exposure-response modeling for lung cancer and mesothelioma. However, the SAB recommends that the agency provide more support for its choice of statistical models for the exposure-response analysis. The SAB also recommends consideration of several models in addition to the Poisson and Cox models used in the draft assessment.

- The agency has been overly constrained by reliance on model fit statistics as the primary criterion for model selection. The SAB recommends graphical display of the fit to the data for both the main models and for a broader range of models in the draft document to provide a more complete and transparent view of model fit. The SAB also recommends that the EPA consider literature on epidemiological studies of other amphiboles for model selection for dose-response assessment, since the size of the Libby subcohort used in the exposure-response modeling is small.

- The EPA has summarized many sources of uncertainty, sometimes quantitatively, as well as the direction and magnitude of the likely impact of each source of uncertainty. The SAB recommends that model uncertainty be evaluated by estimating risks using a more complete set of plausible models for the exposure-response relationship. This sensitivity analysis, while not a full uncertainty analysis, would make explicit the implications of these key model choices.
- Finally, the SAB has identified critical research needs for epidemiological studies, mode of action, and measurement methods for LAA to strengthen future LAA assessment.

The SAB appreciates the opportunity to provide the EPA with advice on this important subject. We look forward to receiving the agency's response.

Sincerely,

/signed/

Dr. David T. Allen, Chair
Science Advisory Board

/signed/

Dr. Deborah L. Swackhamer, Immediate Past Chair
Science Advisory Board

/signed/

Dr. Agnes Kane, Chair
SAB Libby Amphibole Asbestos Review Panel

Enclosure

1. EXECUTIVE SUMMARY

EPA's Office of Research and Development (ORD) requested the Science Advisory Board (SAB) to conduct a peer review of EPA's draft Integrated Risk Information System (IRIS) assessment, entitled *Toxicological Review of Libby Amphibole Asbestos (August 2011)*. The draft document is the first IRIS assessment specific to Libby Amphibole asbestos (LAA), a term used to refer to the mixture of amphibole mineral fibers identified in the Rainy Creek complex near Libby, Montana. The SAB was asked to comment on the scientific soundness of the hazard and dose-response assessment of LAA-induced cancer and non-cancer health effects (see Appendix A).

The SAB finds the EPA's draft assessment to be comprehensive and generally clear, logical and well-written. There are several areas that need more consideration, and we provide recommendations to further enhance the clarity and strengthen the scientific basis of the analyses. The SAB's major findings and recommendations are summarized below.

Mineralogy

The SAB notes that the section on mineralogy provides an important foundation for understanding the properties of Libby Amphibole asbestos (LAA) as related to the evaluation of its potential toxicity and carcinogenicity. The SAB recognizes that physical-chemical characteristics of asbestos (e.g., mineral composition, fiber dimensions) have not typically been available in toxicity studies of LAA. The SAB encourages a more rigorous and accurate description of LAA in the document, while acknowledging the potential ambiguities in the use of mineral-species names in toxicity studies.

Fiber Toxicokinetics

The SAB finds the section on fiber toxicokinetics does not distinguish between chrysotile and amphibole fibers. Since the focus of the draft document is on LAA fibers, it would be better to limit most of the literature reviews and discussion to those dealing with the family of amphibole asbestos fibers. The authors of this section should draw on more authoritative and comprehensive reviews in the literature to correctly specify and clarify issues on deposition and dosimetry.

Noncancer Health Effect

Selection of Critical Studies and Effects

The SAB supports the EPA's selection of the Marysville, Ohio, cohort for development of the RfC. The SAB finds it reasonable to select the subcohort for the main analysis (118 workers who began work in 1972 or later when exposure data were available and who had X-rays from the 2002-2005 exam), with the full cohort of 434 workers used for additional substantiating analysis. However, the SAB recommends additional analyses/cohorts to strengthen and support the RfC since the size of the Marysville subcohort is small. In addition to localized pleural thickening (LPT), the SAB suggests that the EPA consider any X-ray abnormalities as the outcome: LPT, diffuse pleural thickening (DPT), or asbestosis. The SAB also suggests that the EPA conduct analogous analyses (to the extent the data permit) of pleural abnormalities among the Libby workers cohort and the Minneapolis Exfoliation Community cohort.

The SAB agrees that the radiographic evidence of LPT in humans is the appropriate adverse critical effect for the derivation of the RfC. LPT has the appropriate specificity and is not confounded by cigarette smoking. It is a permanent structural, pathological alteration of the pleura and is generally associated with reduced lung function. The reported findings are compatible with the animal data showing tissue injury and inflammation. The SAB has identified additional relevant publications and recommends that the agency include a more detailed review of the literature to further support this conclusion.

Use of Animal and Mechanistic Studies

In general, the SAB finds the laboratory animal studies identified in Tables 4-15 and 4-16 and summarized in Appendix D of the EPA draft report to be appropriate and complete. Laboratory animal studies using a variety of non-inhalation routes of exposure have been used to ascertain the potential fibrogenic and carcinogenic potential of LAA. While inhalation is regarded as the most physiologically relevant means of fiber exposure in animals, there is no published study using this route of exposure for delivery of LAA to experimental animals. Therefore, the deposition and clearance of LAA has not been adequately assessed in experimental animals. However, inhalation studies have been conducted with tremolite, an asbestiform amphibole that is a component of LAA. The potency of inhaled LAA from epidemiology studies should be compared with that of tremolite fibers in rodents to add new information for refining the RfC for LAA.

Carcinogenicity

Weight of Evidence Characterization

The SAB supports the EPA's conclusion that the weight of evidence for LAA is "Carcinogenic to Humans by the Inhalation Route," in accordance with EPA's *Guidelines for Carcinogen Risk Assessment.* The occupational studies showed dose-related increased risks of lung cancer and mesothelioma among workers exposed by inhalation. Effects from short-term intra-tracheal instillation studies in mice and rats include altered gene expression, collagen induction, and inflammatory responses, and are consistent with the early-stage pathological change induced by other amphibole fibers. The EPA also has provided supporting evidence of the carcinogenic potential of LAA from studies with tremolite fibers, in light of LAA being about 6 percent tremolite by composition.

Mode of Action

The SAB finds the weight of evidence for the mode of action (MOA) of LAA based on laboratory studies to be weak. However, there are abundant MOA data for other amphiboles such as crocidolite and tremolite that are likely similar to the MOA for LAA. The SAB views the mode of action of LAA as complex, and recommends that a formal mode of action analysis of LAA be conducted in accordance with EPA's *Guidelines for Carcinogen Risk Assessment.* Based on this formal analysis, the agency may still conclude that the use of the default linear extrapolation at low doses is appropriate.

Selection of Critical Study and Endpoint

The SAB concludes that the EPA's selection of the Libby cohort for the derivation of the inhalation unit risk (IUR) is scientifically supported and clearly described. This cohort has been studied thoroughly,

with detailed work histories and a job exposure matrix. This cohort had elevated asbestos exposure, a wide range of measurements of asbestos exposure, and available cancer mortality data.

The SAB finds the use of the subcohort post-1959 may be reasonable due to the lack of exposure information in many of the workers in earlier years; out of 991 workers hired before 1960, 706 had all department and job assignments listed as unknown.

The SAB supports the use of lung cancer and mesothelioma as endpoints for derivation of the IUR. Since determining the cancer outcome from mortality rather than incidence data may have resulted in an undercount of both cancer outcomes, the SAB recommends more detailed discussion on how the use of mortality data could impact the derived IUR. It also would have been useful to know other major categories of mortality in this cohort.

Use of Laboratory Animal and Mechanistic Studies

The SAB agrees that the database of laboratory animal and mechanistic studies pertaining to LAA is appropriately presented in the report and its Appendices for support of its analysis of the human effects observed. However, the SAB finds the body of the document deficient in not utilizing what is known about the dimensions of the administered fibers from Appendix D. It is generally accepted that differences in biological potency among the various amphibole fiber types are due primarily to differences in dimensions, especially in fiber length distributions. The SAB also recommends that Section 4.6.2.2 be modified to reflect that there are insufficient data to determine the mode of action for LAA.

Inhalation Reference Concentration (RfC)

Estimates of Human Exposure Concentration

The approach described (in Appendix F of the EPA document) for exposure reconstruction is detailed and specific. Due to large uncertainties associated with the unmeasured pre-1972 exposures, the SAB agrees that the draft document appropriately eliminates this set of estimates and adheres only to exposure estimates based on measured concentrations for the derivation of the RfC.

With regard to the exposure metric, the SAB recommends that the EPA re-evaluate the raw exposure data and review pertinent sampling documentation to bolster its use of the geometric mean to represent the job group exposures, rather than an estimate of the arithmetic mean. The agency should consider whether a sensitivity analysis using the minimum variance unbiased estimator (MVUE) of the mean is warranted in the development of the cumulative exposure metric.

Exposure-Response Modeling

EPA's approach to the primary exposure-response modeling was generally appropriate, but the SAB recommends that the procedure be refined and the document should provide a clearer description of how the "best" model was chosen, in accordance with EPA's 2012 *Benchmark Dose Technical Guidance.* Since the Marysville cohort does not support precise estimation of the plateau, the EPA should consider fixing the plateau level based on a study of highly exposed asbestos insulation workers.

The SAB suggests examining other exposure metrics besides the simple cumulative exposure, such as time-weighting of exposures. In addition, the document uses a 10% Extra Risk (ER) as the benchmark response level (BMR) which is not typically used for human quantal response data. The SAB recommends that EPA explain what features of the dataset or outcome variable led the agency to choose a BMR that is considerably greater than the norm for epidemiological data.

Alternative Modeling Approach

The SAB agrees that the rationale for performing additional analyses of the full Marysville cohort is scientifically justified; the analysis of the entire cohort increases the number of cases of LPT available for analysis and substantiates the RfC estimated using the subcohort. However, the SAB recommends that the EPA revise its modeling approach and remove "time since first exposure" (TSFE) from the model of the plateau. EPA should determine whether it is appropriate to use TSFE in the linear predictor alongside cumulative exposure and/or use an alternative exposure metric that incorporates TSFE. The SAB also recommends the revised procedures for the subcohort analysis be followed, such as fixing the plateau using literature values.

Evaluation of Potential Confounders and Covariates

The SAB recommends a revised strategy for evaluation of confounders and covariates. Since the quantity of interest in the analyses of the Marysville cohort is the point of departure (POD), the evaluation of the various covariates should be made with respect to this quantity. The SAB suggests that the covariates fall into two classes: *exposure-related covariates* (various exposure metrics and TSFE) and *non-exposure-related covariates* [age, body mass index (BMI), gender, and smoking status]. For non-exposure related covariates, no additional primary analyses are needed. For exposure-related covariates, the SAB recommends that additional work be done to refine the models to consider alternative exposure metrics, as well as the inclusion of TSFE or other time-related variables in the analyses of the full cohort.

Conversion from Cumulative Occupational Exposure to Lifetime Exposure

The modeled POD is based on cumulative exposure estimates for the worker cohort examined. The SAB recommends using the full 70-year lifetime when converting cumulative to continuous exposure rather than 60 (70 minus the lag of 10 used for exposure in the POD derivation); i.e., do not correct for the lag of 10 for a 10-year lagged exposure, since the time of disease onset is not known in prevalence data.

Selection of Uncertainty Factors

The uncertainty factors deserve additional consideration and analysis. A composite uncertainty factor of 100 (an intraspecies uncertainty factor of 10 to account for human variability and sensitive subpopulations; and a database uncertainty factor of 10 to account for database deficiencies) was applied to the POD for derivation of the RfC. Although it may be difficult to identify specific data on LAA to support departure from the default value of 10 for human variability, concern for the impact on susceptible subpopulations, especially women and children, remains an issue. Consideration of additional data (Minnesota cohort and data on other amphiboles) might support a lower value, such as 3, for UF_D. In addition, a subchronic-to-chronic uncertainty factor higher than 1 may be used, given that the mean and maximum exposure duration in the study are well below the lifetime exposure of interest.

There also is concern that the BMR of 10% for a severe endpoint is not reflected by the choice of a LOAEL- to-NOAEL uncertainty factor (UF_L) of 1.

Characterization of Uncertainties

Overall, the SAB found that while the discussion on uncertainties in the methodology and approach on the derivation of the RfC was thorough, detailed, and logical, the uncertainty assessment can be strengthened. The SAB recommends that additional work be done to substantiate the RfC estimate through additional sensitivity analyses and discussion of results and insights from other datasets and studies.

Inhalation Unit Risk (IUR)

Exposure-Response Modeling

The SAB supports the agency's reliance on the Libby worker subcohort for derivation of the IUR because of its focus on good quality exposure data that are specific for LAA. However, it is important to acknowledge that this small subcohort may have its own limitations as a basis for modeling exposure-response relationship that might be expected in a larger population exposed over a lifetime. The SAB had particular concern about adequate characterization of early life exposures and the potential time dependence for development of disease.

The SAB agrees that the agency clearly described the methods used to conduct the exposure-response modeling for lung cancer and mesothelioma. However, given limitations in the subcohort and other statistical considerations, the SAB made a number of recommendations for providing greater support for this choice of modeling approach and for characterizing model uncertainty.

Having made these points, the SAB recognizes that the agency did conduct extensive sensitivity analyses of their chosen models in various ways to characterize exposure in the Libby cohort. However, the analyses rely on essentially the same underlying models. They do not address the fundamental question of model uncertainty – that is, whether any one model can or should be assumed to represent the exposure-response relationship for LAA. This issue is of particular concern for the estimation of risks from partial lifetime exposure where risk is essentially assumed to be independent of when in the course of a lifetime exposure occurs. Recommendations for addressing model uncertainty are discussed under response to charge question 5 in Section 3.2.6.5.

Approach for Quantification of Inhalation Unit Risk

In order to derive an IUR that represents the combined risk of mortality from lung cancer and mesothelioma, a cancer-specific unit risk for each tumor type was calculated according to the *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005) by linear extrapolation from the corresponding POD. The IUR was then determined as a combined upper bound risk estimate for mortality considering both cancers. The SAB considers the approach to be consistent with the agency's own guidance, and found the description of the procedure used to be clear. However, the SAB recommends that EPA acknowledge that the assumption of independence is a theoretical limitation of the analysis and should provide a fuller justification for this assumption.

Potential Confounding by Smoking

The SAB agrees that the agency's use of the Richardson (2010) method for exploring possible confounding for smoking was appropriate. However, the SAB finds the statement that there is no evidence of confounding by smoking is too strong, and suggests modifications to the discussion that would be more compelling.

Adjustment for Mesothelioma Mortality Under-ascertainment

The number of mesothelioma deaths was adjusted for under-ascertainment stemming from inadequate coding in death certificates. The procedure is not described in any detail, but can be found in Kopylev et al. (2011). The EPA method appears to be scientifically supported, but is not clearly described. The SAB recommends that this section be expanded to provide a more detailed statement of how the numbers were calculated.

Characterization of Uncertainties

The SAB commented that the EPA has summarized the many sources of uncertainty and has evaluated qualitatively, and sometimes quantitatively, the direction and likely magnitude of their impact on uncertainty in the IUR. However, the SAB notes that an important source of uncertainty, that of model uncertainty, might not be accounted for either in the sensitivity analyses conducted to date or in the use of the 95% upper confidence limit (UCL). The SAB recommends that a more straightforward and transparent treatment of model uncertainty would be to estimate risks using a more complete set of plausible models for the exposure-response relationship. This sensitivity analysis would make more explicit the implications of these key model choices for uncertainty in the IUR.

Long-Term Research Needs

The SAB identifies long-term research needs for epidemiological studies, mode of action, and measurement methods for LAA.

- The National Institute for Occupational Safety and Health (NIOSH) and Agency for Toxic Substances and Disease Registry (ATSDR) should continue to monitor mortality among Libby workers and residents of Libby and Troy.
- The SAB recommends future research on mode of action on LAA to focus on biomarkers that are more clearly and specifically related to non-cancer endpoints (i.e., asbestosis) or cancer endpoints (e.g., mesothelioma). Inhalation studies in animal models that can provide both quantitative as well as mechanistic insight should be included.
- EPA should develop a TEM method that provides equivalent data to PCM for LAA.

3.2.3. Noncancer Health Effects of Libby Amphibole Asbestos

3.2.3.1. Selection of Critical Studies and Effects

Question 1. An occupational cohort of workers in a Marysville, OH facility exposed to Libby Amphibole asbestos (Lockey et al., 1984; Rohs et al., 2008) was selected as the basis for the derivation of the reference concentration (RfC). Please comment on whether the selection of this study population is scientifically supported and clearly described. If a different study population is recommended as the basis for the RfC, please identify this study and provide scientific support for this choice.

The rationale for the use of the Marysville, Ohio, cohort for development of the RfC was well described and scientifically supported. However, there are clear drawbacks to this cohort due to the lack of exposure sampling prior to 1972 when most of the cohort began work, the use of self-reported work histories, the end of Libby vermiculite use in 1980 and the mixture of vermiculite sources used throughout the life of the plant. These drawbacks are offset by the solely occupational exposure of this cohort, the use of better quality radiographs taken for research purposes, the use of 2000 ILO standards for reading radiographs, and a cohort with exposures closer to environmental levels. The selection of the subcohort for the main analysis has a clear and strong rationale. (There were 118 workers who began work in 1972 or later when exposure data were available, and who had X-rays from the 2002-2005 exam.) The full cohort of 434 workers was used for analyses to substantiate the subcohort findings.

Although the SAB agrees that the Marysville subcohort represents the best population upon which to base the RfC, there was discussion about the need for additional analyses/cohorts to strengthen and support the RfC since the size of the Marysville subcohort was small. One suggestion is to use the Marysville cohort but include any X-ray abnormalities as the outcome [LPT, diffuse pleural thickening (DPT), or asbestosis]. In addition, cause of death might be assessed for those who died between the two exams. Another suggestion for providing support and perspective to the Marysville findings is to conduct analogous analyses (to the extent the data permit) of pleural abnormalities among the Libby workers cohort (Larson et al., 2012) and among the Minneapolis exfoliation community cohort (Adgate et al., 2011; Alexander et al., 2012). The Libby workers have higher, well characterized occupational exposures compared to the Marysville cohort. The Minneapolis cohort of non-workers generally had estimated exposures at the lower end of the Marysville cohort but included women and children, thus providing a cohort more representative of the general population. However, because the Minneapolis cohort had estimated, not measured exposures, it would not be suitable for the primary RfC analysis. Similarly, because the Libby workers have both environmental and occupational exposures, this cohort should not be used for primary RfC analysis.

Question 2. Radiographic evidence of localized pleural thickening in humans was concluded by EPA to be an adverse effect and was selected as the critical effect for the derivation of the RfC. Pleural thickening is associated with restrictive lung function, breathlessness during exercise and, for some individuals, chronic chest pain. Please comment on whether the selection of this critical effect and its characterization is scientifically supported and clearly described. If a different health endpoint is recommended as the critical effect for deriving the RfC, please identify this effect and provide scientific support for this choice.

Radiographic evidence of localized pleural thickening (LPT) in humans is the appropriate adverse and critical effect for the derivation of the RfC. This is clearly described and well supported by the lines of evidence presented in section 4.1.1.4.2. However, the SAB believes additional evidence is available to further support this view and should be reported.

While other health endpoints (such as diffuse pleural thickening and small opacity profusion) might have been considered candidates for the critical effect for deriving the RfC, the use of LPT is appropriate and well supported. LPT is a permanent, structural, pathological alteration of the pleura. LPT is found at a significantly elevated prevalence in exposed individuals, has the appropriate specificity and is not confounded by cigarette smoking. LPT also is associated with reduced lung function. Furthermore, the findings reported in this section are compatible with the animal data showing tissue injury and inflammation.

It is important to provide a more detailed review of the literature to support the use of LPT as the appropriate endpoint, including studies addressing the relationship between LPT and both pathologic and physiologic abnormalities. Published studies that address the relationship between LPT and lung function suggested by the SAB include Lilis et al., 1991b; Paris et al., 2009; Clin et al., 2011; Sichletidis et al., 2006; Whitehouse, 2004; and Wilken et al., 2011, along with those referenced in the American Thoracic Society (ATS) Statement entitled, *Diagnosis and Initial Management of Nonmalignant Diseases Related to Asbestos: Official Statement of the American Thoracic Society* (ATS, 2004) (Ohlson et al., 1984; 1985; Jarvolm and Sanden, 1986; Hjortsberg et al., 1988; Oliver et al., 1988; Bourbeau et al., 1990; Schwartz et al., 1990; Miller et al., 1992; Van Cleemput et al., 2001; Miller, 2002;). Consistent with that ATS Statement, the SAB concludes that cohort studies have shown significant reduction in lung function, including diminished diffusing capacity and vital capacity associated with LPT. To help clarify the difference between "clinically significant" effects of plaques in a given patient vs. epidemiological studies evaluating the effects of asbestos exposure in an exposed population, the SAB suggests that the EPA clarify in the assessment the range of endpoints that generally can be used to derive an RfC.

In addition to localized pleural thickening, the SAB also suggests that the EPA consider looking at LPT, DPT and small opacity profusion score together as an outcome. There is evidence that LPT is not always the first adverse effect that is detected on chest radiographs, and some individuals with LAA exposure can develop either DPT or increased profusion of small opacities without developing evidence of LPT. Combining outcomes is appropriate, since DPT and small opacity profusion also are effects of asbestos exposure and the goal is to define an exposure level below which LAA is unlikely to have adverse health effects.

Recommendations:

- The SAB suggests the EPA assessment clarify the range of endpoints that generally can be used to derive an RfC.
- The agency should include a more detailed review of the literature to support the selection of LPT through detailing the studies that show the relationship between LPT and both pathologic and physiologic abnormalities, and also risk of other non-cancer asbestos-related diseases.
- In addition to LPT, the document should include an analysis that uses all radiographic outcomes (LPT, DPT and small opacities), recognizing this change may have little impact on the current analysis.

3.2.5.3. Alternative Modeling Approaches

Question 3. EPA's assessment also provides the results of alternative modeling approaches to derive a POD for localized pleural thickening. This modeling used the full Marysville worker data set with exposures from 1957 and later and a Cumulative Normal Michaelis-Menten model that incorporates both cumulative exposure and time from first exposure as explanatory variables. Please comment on whether EPA's rationale for presenting these alternative approaches is scientifically justified and clearly described. Please identify and provide the rationale if a different approach for identifying the most appropriate population within the cohort of Marysville workers is recommended as the basis for estimating a POD.

The SAB notes that this question applies to the full Marysville cohort. The SAB agrees that the rationale for performing additional analyses of the full Marysville cohort is scientifically justified and that the analysis of the entire cohort increases the number of cases of LPT available for analysis and substantiates the primary RfC estimate derived from the subcohort.

However, the SAB does not find the rationale for the analysis approach to be well justified and it recommends that the full cohort analysis be redone. With respect to the approach:

- It is not clear that the scientific basis of using time since first exposure (TSFE) is well founded. EPA should consider what TSFE is supposed to be measuring and how it is related to other variables in the dataset (specifically age and exposure). There is some suggestion in the draft document that in this dataset it is a surrogate measure of intensity since people with larger TSFEs would be more likely to have been exposed to higher levels of LAA present during the early time periods. This perspective should help identify modeling options.

- The SAB also finds that the method for incorporating TSFE into the full cohort analysis is not well justified. Currently, the EPA uses TSFE as a predictor for the plateau in the Cumulative Normal Michaelis-Menten model. No biological justification is given for why this maximum proportion would vary with TSFE.

Regarding revisions to the analysis, the SAB recommends that in this dataset a more natural way to incorporate TSFE into the model would be to allow TSFE to affect the rate of change in the probability of LPT by: (1) including it directly in the linear predictor portion of the model alongside cumulative exposure; and/or (2) using an alternative exposure metric such as residence time weighting (RTW) that more heavily weights exposure in the distant past. The functional form of TSFE could then be selected using standard approaches (e.g., comparing AICs). Since adding TSFE to the model should affect the coefficient of cumulative exposure, the EPA should consider a dichotomous Hill model which allows an exposure parameter (b in Table 5-4) to be estimated, as an alternative to the Michaelis-Menten model. Finally, the SAB recommends that other changes to the analysis follow the approaches used for the subcohort analysis, such as fixing the plateau using literature values as recommended in the response to charge question 2 in Section 3.2.5.2 of this report.

The SAB notes that in principle it may be preferable to base the RfC on an analysis of incidence rather than prevalence data. Because of the nature of the dataset, the Marysville cohort does not support a direct analysis of incidence. While it may be possible to fit an alternative model derived from integration of a plausible incidence model (e.g., see Berry et al., 1979; Berry and Lewinsohn, 1979; Paris et al., 2008), this approach will require a number of untestable assumptions, particularly given the small size

of the Marysville cohort. In lieu of conducting such an analysis, the SAB recommends that an explicit acknowledgement be added to the report regarding the implications of various model alternatives.

Recommendations:

- Improve the scientific justification for using TSFE in the full cohort analysis; this justification will include an explanation of its meaning in the context of this dataset.
- Revise the full cohort analysis to change the approach to incorporating TSFE, removing it from the model of the plateau. As part of the revision, the SAB suggests assessments be made to determine whether it is appropriate to use (a) the dichotomous Hill model, (b) TSFE in the linear predictor alongside cumulative exposure and/or use an alternative exposure metric that explicitly incorporates TSFE, and (c) the approaches recommended for the subcohort such as a fixed plateau. As appropriate, such analyses should include assessment of the functional form of TSFE.
- The SAB recommends that the EPA present the lower 95% confidence limit of the benchmark concentration (BMCL) estimates from a set of reasonable and plausible models, and selections of data, which will both inform selection of a preferred model and illustrate the range of model uncertainty.

3.2.5.4. Potential Confounders and Covariates

Question 4. EPA has evaluated potential confounders and covariates where data are available. Specifically, EPA has explored the influence of age, body mass index, smoking status, time since first exposure, gender, and alternative exposure metrics on model fit and evaluated their association with the modeled health outcomes (see Section 5.3). Are these analyses clearly described and appropriately conducted? Are the results of these analyses appropriately considered in the RfC derivation? Additionally, there is a possibility of exposure-dependent censoring in participant selection for the update of the Marysville cohort (Rohs et al., 2008) but no evidence of selection bias. Does the SAB have any specific recommendations for evaluating and, if appropriate, quantitatively addressing exposure-dependent censoring in these analyses?

The SAB recommends a revised strategy for evaluation of covariates. The target of inference for the analyses of the Marysville cohort is the POD, which in this case is the BMCL. The evaluation of the various covariates should be made with respect to this target of inference. The SAB suggests the covariates fall into two classes: *exposure-related covariates* (various exposure metrics and TSFE) and *non-exposure-related covariates* [age, body mass index (BMI), gender, and smoking status]. We provide recommended revised strategies for considering these two classes of covariates that follow directly from consideration of the target of inference.

Non-exposure-related covariates: A decision on whether to control for the non-exposure-related covariates should account for how the EPA wishes to determine and apply the RfC. The SAB suggests a BMCL most directly applicable to all members of the general population is most appropriate. This implies that the BMCL should be estimated from a model that includes exposure covariate(s), but that is otherwise unadjusted. This is the same approach used in the current draft document; only the rationale for the approach is different. The SAB suggests it would be informative to conduct sensitivity analyses to examine how the BMCL varies across subgroups defined by covariate values (e.g., older males or smokers). Because the Marysville subcohort is a small dataset, it is difficult to conduct this evaluation exclusively in the subcohort. Therefore the SAB suggests that the EPA use the *full* cohort for the model selection and parameter estimation components of sensitivity analyses incorporating these covariates.

For this activity the EPA would use its selected final model after excluding all exposure variables (e.g., the dichotomous Hill model with fixed background, fixed plateau, and after dropping exposure variables). After fitting a model with a specific set of non-exposure-related covariates in the full cohort, one can estimate a "risk score" (i.e., the linear predictor for the non-exposure-related covariates). This risk score would be included as a single term (as either an unscaled offset or scaled by its estimated coefficient) in the subcohort analysis. Similar to the approach presented in Table E-5, these analyses can be used to produce a new table of subgroup-specific conditional BMCLs; these values will give some evidence of how the target of inference varies by subgroup. In addition, weighted averages of the conditional BMCLs can be computed to reflect population average BMCLs for specific covariate distributions in target populations. For instance, Gaylor et al. (1998) gives a formula for the upper tail of a 95% confidence interval and this formula can be extended to obtain BMCLs for weighted averages.

Exposure-related covariates: The inclusion of exposure-related covariates in the model is fundamental to the inference. The EPA has done excellent preliminary work, and the SAB has provided recommendations in Sections 3.2.5.2 and 3.2.5.3 of this report about how to revise the approach. In addition the SAB recommends that the EPA consider taking several further steps. First, alternative exposure metrics should be assessed directly in the subcohort dataset to determine whether they fit the data better. In particular, alternative metrics (such as residence time weighted exposure) that more heavily weight more distant exposure may be more biologically plausible because individuals exposed at an earlier age might be more susceptible to the damaging effects of asbestos. Second, TSFE should be considered for addition to the model. Since TSFE is complete and equally well estimated across all members of the cohort, the full cohort can be used to determine how to model this variable. Similar to the approach recommended for the sensitivity analyses discussed above, this would be done using the model intended for the subcohort, but omitting exposure variables other than TSFE. Then, the functional form of TSFE selected using the full cohort can be added to the subcohort analysis, either as an unscaled offset term or as a scaled covariate. Given biological understanding of the disease process, for models with both estimated exposure and TSFE included, it would be appropriate to report the BMCL conditional on a large TSFE.

Additional comments on covariates:

- BMI: In section 5.2.3.3.1., it would be helpful if the justification for considering BMI as a covariate were briefly explained. It is included elsewhere, but readers may have missed it.
- TSFE:
    - TSFE deserves careful consideration for both biological and dataset-specific reasons. It is an important determinant of LPT both because individuals' lung tissues exposed at an earlier age might be more susceptible to the damaging effects of asbestos and because asbestos' effect over time is increasingly damaging. It is correlated with exposure in this dataset since subjects with the longest TSFE were exposed in the early years of the cohort when exposures were higher. It is also more accurately estimated than exposure.
    - The SAB does not agree with the use of the Cumulative Normal Michaelis–Menten model to adjust for TSFE because it makes the assumption that the TSFE only affects the plateau. This has not been justified biologically or in the context of features of this particular dataset. Instead, the SAB recommends that EPA consider alternative approaches to account for TSFE.

- Smoking:
  - Smoking is included in the follow-up by Rohs et al. (2008). However, the ever/never categorization of smoking is much less informative than the pack-year analysis of smoking used in the earlier study by Lockey et al. (1984).
  - There is an important discussion of the evidence linking pleural changes and smoking in footnote 34 on page 5-46. This information could be moved into the body of the report, and amplified somewhat. A table summarizing the relevant studies (irrespective of type of amphibole asbestos) summarizing the evidence regarding the role of smoking would be useful.
- Gender: There is little discussion of gender, except in places where the number of females is listed as too few to analyze in any detail. The SAB did not regard this as a serious concern because it is reasonable to assume that females and males have similar probabilities of developing LPT.

The SAB recommends that a table be included summarizing the results of the various sensitivity analyses and how they change the POD.

Exposure-dependent censoring: The exposure-dependent censoring discussion is based on results from Rohs et al. (2008) that inappropriately separated deceased non-participants from the remaining non-participants. Once all non-participants are combined there is no evidence of exposure-dependent censoring. Furthermore, exposure-dependent sampling by itself does not lead to bias in risk estimates. The important issue for bias is whether two individuals with the same exposure, one diseased and the other not, are equally likely to participate in screening. There has been no strong rationale presented that would indicate that such differential selection has occurred in this cohort.

Recommendations:

- Revise consideration of covariates to focus on their impact on the target of inference.
  - For non-exposure-related covariates, this only alters the presentation; no additional primary analyses are needed. Sensitivity analyses conditional on subgroups defined by covariates can be added.
  - For exposure-related covariates, additional work is needed to refine the models to consider alternative exposure metrics, as well as the inclusion of TSFE or other time-related variables in analyses of the full cohort. The SAB encourages the EPA to either fully justify analyses based on the Cumulative Normal Michaelis-Menten model in the context of this particular dataset, or replace them.
- Revise this discussion of Rohs et al. (2008) to make note (perhaps in a revised table) that the dose distribution in participants is similar to the overall dose distribution of the original full cohort. Furthermore, revise the discussion of exposure dependent sampling to distinguish this from bias differential sampling in the sense above.

With respect to exposure assessment, analytical methods and environmental conditions are substantial contributors to uncertainty because of differences between the 1970s and today. As discussed throughout the report, PCM was the only generally accepted method for measuring airborne fiber concentrations used until the 1980's. PCM's limitations are well-detailed in the report: an inability to detect fibers smaller than 0.25 μm, an inability to differentiate asbestos fibers from other fibers, and a limitation to counting only fibers longer than 5 μm. Today, TEM can easily detect and positively identify airborne asbestos of all sizes. But, because the RfC is based on 1970's PCM analyses, the RfC must be implemented in a way that most closely replicates analysis in the 1970's. At the 1970's study site, the vast majority of measured fibers were almost certainly LAA, so PCM's inability to identify asbestos did not create much uncertainty. Today, even ambient air will yield fiber concentrations that exceed the RfC. The culprit fibers will likely be cellulose fibers from cotton, wood, paper or synthetic fibers, rather than asbestos. Hence, today's PCM counts will be from fibers that are unrelated to the RfC. Thus it is important that TEM be used to identify and count asbestos fibers in air samples for RfC purposes. Finally, Page 5-118, Lines 22-33 of the EPA's draft document discuss the two-fold under-reporting of fibers because of PCM's poorer resolution in the 1970's, 0.44 μm versus 0.25 μm today. Because today's PCM analysts have no capability for discriminating fibers > 0.44 μm, the need for TEM analysis of samples collected for implementation of the RfC is even more important. A TEM protocol for PCM equivalent fibers wider than 0.44 μm could be easily developed.

Recommendations:

- Harmonize the uncertainty discussions across the document.
- Substantiate the RfC estimate through
    - Additional sensitivity analyses of the subcohort;
    - Discussion of results from other studies;
    - Additional sensitivity analysis of the full cohort; and
    - Summarizing in tabular form the results of the various sensitivity analyses and model alternatives, to show how they affect the POD.
- Use TEM to identify and count asbestos fibers longer than 5, 10, and 20 μm in air monitoring samples for implementation of the RfC.

3.2.6. Inhalation Unit Risk (IUR)

3.2.6.1. Exposure-Response Modeling

Question 1. Exposure-response modeling was conducted separately for lung cancer and mesothelioma mortality. The POD estimates for these endpoints are based upon analysis of the subcohort of workers first exposed after 1959 when the exposure data were judged to be better characterized. The exposure-response modeling included consideration of a variety of exposure metrics that varied with time and incorporated different lag and decay parameters. Based on the results of the exposure-response modeling, a life table analysis was used to determine the PODs for each type of cancer for the various exposure metrics. Have the exposure-response modeling and determination of the PODs from life table analysis been appropriately conducted and clearly described? If a different approach to exposure-response analysis is recommended as the basis for estimating the IUR, please identify the recommended methods and provide a rationale for this choice.

In general, the EPA clearly described the methods it had selected to conduct the exposure-response modeling for lung cancer and mesothelioma. The risk calculations in the life tables appeared correct but would benefit from clearer explanations. Some suggestions for clarifications are noted below.

The agency was overly constrained by reliance on model fit as the primary criterion for model selection and the SAB recommends a broader discussion of biological and epidemiological criteria as well. For the mesothelioma data, for example, the Peto model was disregarded due to a poorer fit than the Poisson model. The results for this analysis are not shown and, given the particular interest in this model, should have been. A parametric survival model (e.g., Weibull) could have also been used to obtain estimates of absolute risk. It would also be appropriate to compare the results of the final model against those from fitting a two-stage clonal expansion (TSCE) model. Use of the TSCE model would allow for a more direct evaluation of, and possibly justification for, age-dependency of the IUR. The Richardson (2008) paper provides a publicly available and transparent approach to application of the TSCE. Ultimately, there are many competing models that could have been used instead of the Poisson and Cox models (e.g., parametric survival models, accelerated failure time models, additive models) that could have provided very different estimates of risk, but they were not discussed.

Data exist that suggest that the lifetime risk of developing the mesothelioma increases the earlier in life that exposure is first received. The Peto model (Peto, 1979; Peto et al., 1982) was developed to explain such observations in the empirical data. While the Peto model has been more widely used for risk assessment, most notably in the previous IRIS summary for asbestos, it has also only been formally fitted to data in a limited number of cohorts (HEI-AR, 1991). Ongoing analysis of incidence of mesothelioma appears to be consistent with the exposure-response relationship described in the Peto model. The draft report needs to do a more complete job of justifying why this and other epidemiologic evidence should be excluded as a basis for selection of a plausible model for predicting mesothelioma risk. Chapters 2 and 3, for example, consider toxicological and other evidence developed with exposures to asbestos that are not strictly LAA. The cohorts used in the development of the Nicholson/Peto model and the exposures they experienced should provide information about the time course of the development of disease.

The SAB recognizes that the agency's effort to focus on good quality exposures specific to LAA has led to reliance solely on the Libby worker subcohort. This rationale is understandable, but at the same time, it is important to acknowledge that this small subcohort may have its own limitations as a basis for modeling exposure-response relationships for a larger population over a lifetime. As a sensitivity analysis to evaluate the potential impact of omitting the Libby workers hired before 1959, the SAB recommends analyzing the entire Libby cohort using interval statistics (Nguyen et al., 2012; Manski 2003; *inter alia*) or other traditional approaches for data censoring in predictors (cf. Küchenhoff et al., 2007). It can be misleading to use midpoint substitution (as described in Section 5.4.6.1.2) that assumes poorly measured or missing predictors have some constant value. Interval statistics and traditional censoring approaches to measurement uncertainty would, in essence, replace point values with interval ranges. When the intervals are narrow, as they might be for 21% of the early hires for which jobs titles are available, there might be a good deal of recoverable information present. When the intervals are much wider, there would be accordingly less information. Whatever empirical information may be present, it is worth evaluating whether its inclusion is better than leaving out the data entirely, which in principle amounts to replacing them with intervals that are completely vacuous, from zero to infinity. This approach can produce an interval range for the final outputs, which would provide the explicit quantitative uncertainty statement as recommended by previous National Academy of Sciences reviews.

The SAB recognizes that the agency did conduct sensitivity analyses with several analyses of the Libby cohort data, including those that used different models (Tables 5-20 for lung cancer and 5-21 for mesothelioma). A limitation of these analyses is that they all rely on the assumption that the effect of exposure can be modeled as a function of cumulative dose. This assumption is consistent with the agency's *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005), which state that "unless there is evidence to the contrary in a particular case, the cumulative dose received over a lifetime, expressed as an average daily exposure prorated over a lifetime, is recommended as the appropriate measure of exposure to a carcinogen." EPA therefore did not address the fundamental question about whether any one model can or should be assumed to represent the exposure-response relationship for LAA. Therefore, one cannot be confident that the "true" exposure-response relationship for LAA is really "accounted for" by use of the upper confidence limit (UCL) on the slope (per fiber/cc) or, ultimately, the combined IUR from mesothelioma and lung-cancer mortality (see related discussion in response to question 3 and 5 in Section 3.2.5).

This issue is of particular concern for the estimation of mesothelioma risks from partial lifetime exposures, where risk is essentially assumed to be independent of when in the course of a lifetime exposure occurs. For example, one year of exposure to a given concentration in childhood yields the same lifetime average daily dose as one year of the same exposure in adulthood. This assumption is not consistent with the relevant body of evidence on the development of asbestos-related disease. Therefore, there is some probability — not well characterized — that this approach underestimates the relative effect of early exposure, but exaggerates the effect of exposure later in life.

Recommendations:

Two types of recommendations have been made. The first set is asking for simple explanations in the text that the SAB thinks will clarify the rationale for analytic choices made by the EPA. The next set includes requests for additional presentations of data or analyses, roughly in order of priority, that the SAB concludes are important to provide some quantitative perspective on the analytic choices made.

Clarifications:

- Poisson regression analyses: the mathematical form of the regression function should be given, and discussion of whether the potential for over-dispersion was assessed.
- Cox proportional hazards modeling: the reasons should be given for not conducting a Bayesian analysis as was done for the Poisson regression model for mesothelioma.
- Life-table analysis: the method used to estimate the hazard function for the exposed population should be clearly spelled out in the text. Was it based on a nonparametric estimate of the baseline hazard from the sub-cohort? Given that the SEER data were used to calculate the background incidence of lung cancer, it would seem more appropriate to use those data to estimate the baseline hazard and then to use the regression coefficient obtained from the Cox model applied to the sub-cohort data to obtain the hazard of the exposed group. Thus, the reasons for not using the SEER data to estimate the baseline hazard should be explained.
- Expand the discussion of model selection to explain the reliance on model fit criteria for model selection. In particular, why should the broader epidemiologic evidence on the time course of disease not argue at least for the presentation of more than one statistical model?

Provision of additional data or analysis:

- In a tabular form, summarize the fit results, POD estimates, and IUR estimates from the full range of models considered in order to show the dependence of the IUR estimate on model selection.
- Present the fit to data graphically for both the main models and for a broader range of models, including the Peto model. This step would provide a more thorough and transparent view of fit, particularly in the region of the BMR, than is allowed by examining summary statistical values alone.
- Provide in an appendix the details of the Nicholson/Peto model fit for which the text currently states "data not shown".
- Allow evaluation of the time dependence of disease by providing tabulations of mesothelioma mortality rates and lung cancer SMRs by time since first exposure, duration of exposure and period of first exposure (for both the full and sub-cohorts of Libby workers).
- Evaluate the feasibility of conducting an ancillary analysis of the full Libby data set, including hires before 1959, using interval statistics or other traditional censoring methods (not simple midpoint substitution). At a minimum, discuss the possible quantitative uncertainties associated with using the smaller subcohort.

3.2.6.2. Potential Confounding by Smoking

Question 2. Smoking is a strong independent risk factor for lung cancer and may be an important confounder of the lung cancer mortality analysis. Data on individual smoking habits and history were largely missing and could not be used to control for potential confounding in regression analyses. However, EPA used three approaches to evaluate the confounding issue, including restriction of the cohort and two analytic evaluations of the potential for confounding by smoking (see Section 5.4.3.6.5). Please comment on whether the methods and analyses are clearly presented and scientifically justified. If additional analyses are recommended, please identify the methods and scientific rationale.

The SAB recognizes the challenges in controlling for smoking given the lack of data on smoking histories for the cohort. The agency has taken reasonable steps to identify the potential for confounding using independent approaches. However, statements in the document (on p. 5-96 and again on p. 5-127) that— because the proportional hazards assumption is satisfied in the subcohort— there is no evidence of confounding by smoking, are too strong. Reaching this conclusion requires some strong assumptions, including one that the decline in smoking prevalence observed in the general U.S. population also occurred in the Libby cohort.

The agency's use of the Richardson (2010) method for exploring possible confounding for smoking was appropriate. However, the conclusion that there is no evidence for confounding by smoking relies more heavily on the *p*-values, which are marginally non-significant, than it needs to. More compelling is the observation of a negative association with COPD in their analyses. The fact that the coefficients for exposure in the COPD Cox models were negative is strong evidence against positive confounding; smoking is positively related to COPD risk and thus if positive confounding is occurring, then one would also expect the relationship between asbestos exposure and COPD risk to be positive.

Recommendations:

- The numbers of COPD deaths (n) in the sub-cohort that were the basis for the analysis should be presented in the text.